

सत्यमेव जयते

INDIAN AGRICULTURAL
RESEARCH INSTITUTE, NEW DELHI

I.A.R.I.6.
GIP NLK—H-3 I.A.R.I.—10-5 55—15,000

PROCEEDINGS

OF THE

ROYAL SOCIETY OF LONDON

SERIES A. MATHEMATICAL AND PHYSICAL SCIENCES

VOL 190

LONDON

Printed and published for the Royal Society
By the Cambridge University Press
Bentley House, N W 1

9 September 1947

Printed in Great Britain at the University Press, Cambridge
(Brooke Crutchley, University Printer)
and published by the Cambridge University Press
(Cambridge, and Bentley House, London)

Agents for Canada and India Macmillan

CONTENTS

SERIES A VOLUME 190

No A 1020. 17 June 1947

PAGE

Shock-wave and boundary-layer phenomena near a flat surface By A Fage, F R S and R F Sargent (Plates 1 and 2) . 1

The assignment of the slow-neutron-produced activities of thallium and the dual disintegration of radium E By E Broda and N Feather, F R S . 20

Some improvements in the use of relaxation methods for the solution of ordinary and partial differential equations By L Fox . . . 31

The escape of molecules from a plane surface into a still atmosphere. By K J Brookfield, H D. N Fitzpatrick, J F Jackson, J B. Matthews and E A Moelwyn-Hughes 59

On the propagation of electromagnetic waves through the atmosphere By B. K Banerjea 67

The intermetallic compound phases of the system aluminium-manganese-zinc By G V Raynor and D. W Wakeman (Plates 3 and 4) . 82

X-ray reflexions from dilute solid solutions By K Huang . . . 102

The adsorption of vapours on mercury. III. Polar substances By C Kemball 117

Sublimation in a Wilson chamber. By B M Cwilong . 137

No A 1021 8 July 1947

Studies on silk proteins I The properties and constitution of fibroin The conversion of fibroin into a water-soluble form and its bearing on the phenomenon of denaturation By D Coleman and F O Howitt (Plate 5) 145

The scattering of 4·2 MeV protons by protons By A N May and C F Powell 170

The scattering of 4 2 MeV protons by deuterium, helium and other light elements By H Heitler, A N May and C F Powell 180

The elastic scattering of 6·5 MeV deuterons by deuterium, helium and other light elements By K M Guggenheimer, H Heitler and C F Powell 196

Ozonosphere temperatures under radiation equilibrium By E H Gowan . 219

Night cooling of the ozonosphere By E H Gowan 227

Turbulence and diffusion in the lower atmosphere with particular reference to the lateral effect By D R. Davies 232

The vibration spectrum and molecular configuration of cyclohexane. By D A Ramsay and G. B. B. M Sutherland 245

The crystal structure of sucrose sodium bromide dihydrate By C A. Beevers and W. Cochran 257

Atmospheric tides in the ionosphere II Lunar tidal variations in the *F* region near the magnetic equator By D F. Martyn 273

No A 1022 12 August 1947

PAGE

On the catalytic dehydrogenation of naphthenes I Kinetic study By E. F G. Herington and E K Rideal, F.R S. . . . 289

The catalytic dehydrogenation of naphthenes II. Competition experiments and energies of C-H bonds By E F G Herington and E K. Rideal, F.R S. 309

The accuracy of bond-length estimations obtained by double Fourier series methods from X-ray data By J M. Robertson, F R S and J. G. White . 329

The density increments of proteins By G. S Adair, F R S and M E Adair 341

Solar radiation at radio frequencies and its relation to sunspots By L. L McCready, J L Pawsey and R Payne-Scott . 357

Influence of space charge on thermionic emission velocities By O Klemperer 376

Oxide cathodes The effect of the coating-core interface on conductivity and emission By D A Wright 394

The establishment of the dilatation effect in a field of sound waves By P M. Davidson 418

The control of amplitude in phase-contrast microscopy. By E W Taylor. (Plates 6 and 7) . 422

Cosmic-ray bursts and shower spread under large thicknesses of lead By J W F Juritz and C B O Mohr 426

No A 1023 9 September 1947

The theory of the magneto-resistance effects in metals By E. H Sondheimer and A H Wilson, F R S. . . . 435

A general kinetic theory of liquids. III Dynamical properties By M Born, F R S and H S Green . . . 455

A dynamical model of a crystal structure By Sir Lawrence Bragg, F R S and J F. Nye (Plates 8 to 21) . . . 474

The accuracy of atomic co-ordinates derived from Fourier series in X-ray structure analysis III By A D Booth . . 482

The accuracy of atomic co-ordinates derived from Fourier series in X-ray structure analysis. IV The two-dimensional projection of oxalic acid By A D Booth 490

Local and extensive penetrating cosmic ray showers By D Broadbent and L Jánossy 497

The crystal structure of some molecular complexes of 4-4′-dinitrodiphenyl II. The complexes with benzidine, *NN N′N′*-tetramethylbenzidine, 4 4′-dihydroxy-diphenyl, 4-iododiphenyl and 4-bromodiphenyl By D H. Saunder (Plate 22) 508

Some apparent periodic errors in the crystal lattice of the molecular complexes of 4 4′-dinitrodiphenyl with 4-iodo- and 4-bromodiphenyl. By R W. James and D H Saunder (Plate 23) 518

Decay of vorticity in isotropic turbulence By G. K Batchelor and A A. Townsend 534

Measurements in the turbulent wake of a cylinder. By A A Townsend . 551

The vibration spectrum and molecular configuration of 1:4-dioxane By D A Ramsay 562

Index 575

Shock-wave and boundary-layer phenomena near a flat surface

By A. Fage, F.R.S. and R. F. Sargent

Aerodynamics Division, National Physical Laboratory

(*Received* 21 *September* 1946—*Read* 13 February 1947)

[Plates 1 and 2]

Shock-wave and turbulent boundary-layer phenomena near the smooth flat metal floor of a specially designed supersonic tunnel are studied from traverses made with pitot, static pressure and surface tubes and from direct shadow and Töpler striation photographs. Near-normal and oblique shock-wave systems, with or without a bifurcated foot, are considered.

A near-normal shock wave and the Mach shock wave of an oblique incident and reflected wave system bifurcate when the strength, χ, of the wave above the point of bifurcation attains a definite value. The height of the bifurcated foot and the angle between the two limbs at first increase rapidly with an increase in χ above this value and afterwards much more slowly. The surface pressure rises steeply immediately behind the front limb of a bifurcated foot, whilst the back limb eases the distortion of flow behind. The boundary layer thickens rapidly and the surface friction falls steeply to zero in the space between the two limbs. The loss of stagnation pressure, and so the drag, is smaller than that associated with the flow through the main wave. The rise in surface pressure and the thickening of the boundary layer are less severe, and the surface friction does not fall to zero, behind the foot of a wave which is not bifurcated.

A comparison of the properties of a bifurcated shock wave in an infinite inviscid stream with those measured in the tunnel shows that the boundary layer increases the angles ξ_1, ξ_2 (see figure 3) which the front and back limbs make with the undisturbed stream direction and also the angle of refraction ϵ_f at the front limb. The increases in ξ_1 and ϵ_f are associated with the thickening of the boundary layer behind the front limb, whilst the angle $(\xi_2-\xi_1)$ between the limbs, which is also increased, is dependent on the manner in which the surface layer thickens.

1. Introduction

1·1. A plane shock wave normal to a smooth flat surface aligned in the direction of a uniform stream extends right down to the surface without any change in intensity, when the flow is two-dimensional and the fluid is inviscid.

If the wave impinging on the surface is not normal but oblique, there is a critical angle of incidence to the surface below which the wave is reflected regularly at the surface, the flow beyond the reflected wave being parallel to the surface. If the angle of incidence is greater than the critical angle, the point of intersection of the incident and reflected waves is displaced outwards from the surface and a third wave, the Mach wave, extends from the point of intersection right down to the surface. The fluid which emerges from the reflected wave has a velocity and entropy different from those in the fluid which emerges from the Mach wave, the two fields of flow being separated by a vortex sheet. At the inception of this regime, the Mach wave is normal to the surface, the pressures on the two sides of the vortex sheet are the same, but the density and temperature are different. As the point of intersection of the

incident and reflected waves recedes from the surface the Mach wave becomes more and more curved, but the direction of flow and also the pressure on the two sides of the vortex sheet just beyond the point of intersection tend to be the same.

1·2 For a real fluid, the velocity in front of a shock wave cannot be uniform everywhere, for the velocity in the boundary layer falls to zero at the surface. The conditions near a surface, for both near-normal and oblique wave systems are therefore more complicated than those for an inviscid fluid Clearly, a shock wave cannot extend to the surface itself, for the flow very near the surface is subsonic. Further, if the part of a shock wave in the free stream at some distance from the surface were everywhere normal to the surface, the part near the surface could not be, for the distortion of the local flow arising from the velocity gradient there will be associated with a distortion of the shock wave itself Shock-wave and boundary-layer phenomena near a surface have often been observed and photographed, in particular for airflow through nozzles and past aerofoils and other streamline bodies. The purpose of the present work is to examine such phenomena in some detail for the case of airflow along a smooth flat surface on which the boundary layer is turbulent.

2 Notation

x	longitudinal distance from tunnel entry
y	lateral distance from floor.
h	height of surface tube
h_e	effective height of surface tube
h_f	height of bifurcated foot of shock wave, see figure 3
δ	boundary layer thickness
δ_1	displacement thickness of boundary layer, $\int_0^\delta \left(1 - \frac{\rho u}{RU}\right) dy$.
θ	momentum thickness of boundary layer, $\int_0^\delta \frac{\rho u}{RU}\left(1 - \frac{u}{U}\right) dy$
ρ	air density.
ρ_A	atmospheric density.
R	density just outside boundary layer
ρ_H	stagnation density
ρ_1, ρ_2	density in front of and behind a single-shock wave
p	static pressure
p_A	atmospheric pressure
p_1, p_2	static pressure in front of and behind a single-shock wave.
χ	strength of shock-wave system: $\chi = p_2/p_1$ for a near-normal shock wave and for a Mach shock wave, and p_3/p_1 for an oblique and reflected wave system, see figure 3.
Δ_p	increase in floor pressure measured from toe of shock wave to point where $\tau_0/(\tau_0)_d$ falls to its minimum value

H stagnation pressure of air when brought to rest adiabatically.
H_r stagnation pressure reading of pitot tube.
H_1, H_2 stagnation pressure in front of and behind a single-shock wave
q resultant velocity.
q_1, q_2 resultant velocities in front of and behind a single-shock wave
u velocity in boundary layer
U value of u just outside boundary layer
u_s value of u calculated from surface-tube reading.
u_1, u_2 components of q_1 and q_2 normal to shock wave
v_1, v_2 components of q_1 and q_2 parallel to shock wave
a velocity of sound
a_A atmospheric velocity of sound.
M Mach number, q/a.
a_* velocity of sound for state $M = 1$, $a_*^2 = a_1^2/1\ 2$
β angle between main shock-wave front and undisturbed stream direction
β_1, β_2 values of β for incident and reflected waves respectively, see figure 3*b*
ξ_1 angle between the front limb of bifurcated shock wave and undisturbed stream direction, see figure 3*a*
ξ_2 angle between the back limb of bifurcated shock wave and undisturbed stream direction, see figure 3*a*
α_2 angle between back limb of a bifurcated foot and velocity in front
ϵ angle of refraction caused by a shock wave
ϵ_f value of ϵ for a bifurcated foot of a shock wave
τ_0 intensity of surface friction.
$(\tau_0)_d$ datum value of τ_0 measured 0 2 in. forward of toe of shock wave
ν kinematic viscosity
μ coefficient of viscosity } suffix w denotes wall values.
γ ratio of specific heat at constant pressure to specific heat at constant volume, 1·4 for air.

3 Wind-tunnel apparatus

3·1. The experiments were made in a long 2 in square section tunnel fitted with a faired intake. A diagrammatic sketch of the tunnel is given in figure 1 and a photograph in figure 2. Supersonic conditions in the observation section of the tunnel were obtained by a detachable bulge in the roof of the tunnel, forming a convergent-divergent nozzle Three bulges, designated I, II and III, were used to give $M \approx 1{\cdot}2$, 1 5 and 1 6, respectively, in the observation section. The observation section had vertical walls of plate glass through which flow phenomena could be investigated optically and a flat floor on which boundary-layer phenomena could be observed. Static, pitot and surface tubes were mounted on a sliding flat brass plate, designed to allow them to be traversed up and down stream This sliding floor fitted into longitudinal grooves in the vertical sides of the tunnel and rested on a fixed

wooden floor. The leading edge of the plate was chamfered on the top surface to minimize flow disturbance on the floor The sliding floor was clamped in position by a brass sealing plate bearing on the outer surface of the wooden floor. This plate prevented air leakage into the tunnel, and carried a micrometer screw for traversing the brass floor The flow in the tunnel was induced by a flow of compressed air from an annular injector slot, downstream of the observation section, discharging into a conical diffuser A transition piece between the observation section and the injector slot changed the cross-section shape of the tunnel from square to circular The floor at the entry section of the tunnel was roughened by a line of fine wire spikes projecting about 0 02 in above the surface Further, the values of Ux/ν at the positions in the tunnel where shock-wave and boundary-layer phenomena were studied were large and greater than 6×10^6, so that the flow in the boundary layer at the positions of observation was turbulent.

3·2. The static tube was made of 0·035 in diameter steel tubing with a conical nose, 10° half angle It had four holes 0·01 in diameter, 0·75 in back from the nose and symmetrically spaced around the circumference The tube was supported at its rear end, 2 25 in behind the pressure holes, by a vertical steel tube 0·055 in passing through a hole in the floor The pitot tube was constructed of tubing of the same bore as the static tube The two tubes were mounted side by side, 0·2 in on opposite sides of the central vertical plane of the tunnel, with their pressure orifices in the same cross-section. The pitot tube was designed to have its vertical tube close behind that of the static tube, to minimize tunnel blockage The two tubes could be traversed vertically, side by side, by a micrometer gear A static pressure hole in the floor was used to measure surface pressure Pressures were read on water and mercury manometers

A surface tube, figure 14*b*, was used to explore the flow very near the floor. The tube was offset 0 2 in from the centre line of the floor and the pressure at its mouth was measured against the static pressure taken at a small hole in the floor, situated in the transverse cross-section passing through the tube mouth but offset 0·2 in on the opposite side of the centre line. The Appendix gives a description of the surface tube and of the method of calibration.

4 Shock waves. Photography

4 1 The investigation was made for two types of wave systems· (*a*) near-normal single-shock wave, with or without a bifurcated foot, (*b*) oblique incident and reflected wave system, with or without a bifurcated foot (figure 3). A near-normal shock wave is defined as one for which the part in the free stream at some distance from the floor tends to be normal to the floor The near-normal shock waves investigated are waves terminating the supersonic field in the divergent tunnel beyond the throat The strength of such a wave depends on the Mach number immediately in front of the wave and therefore on its position. The position depends on the power absorbed by the induced air mixing with the injected air. For low blowing

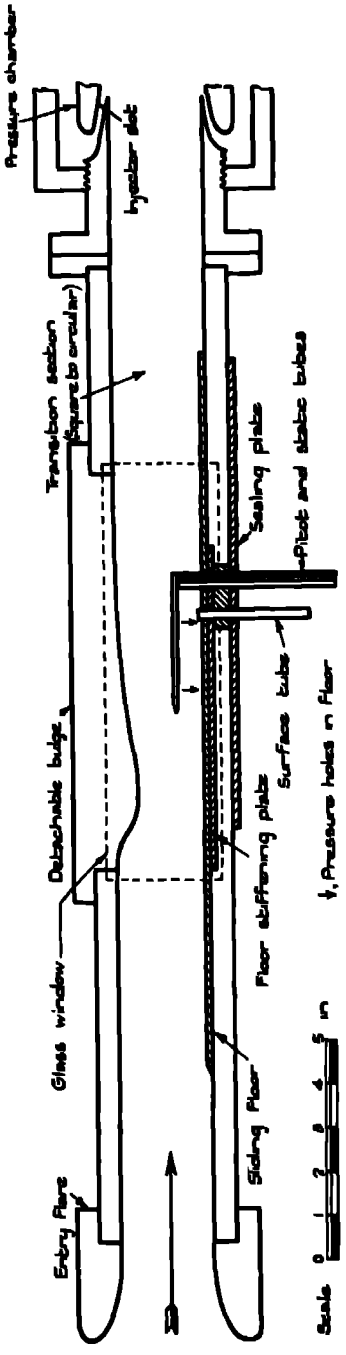


FIGURE 1 Sketch of tunnel

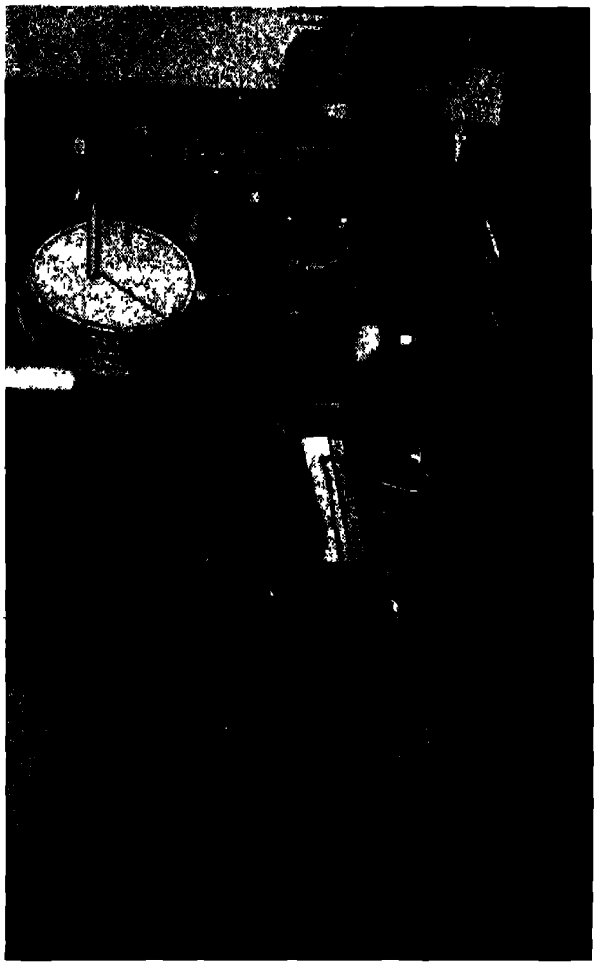


FIGURE 2. Photograph of tunnel and apparatus

pressures of the injected air, the power absorbed was unsteady and caused a shock wave to oscillate rapidly with an amplitude about 1 in on either side of its mean position. With an increase in the blowing pressure the mean position of the shock wave moved downstream until a choke, caused by a displacement due to the injected air, was formed in the induced air downstream of the slot The shock wave then became steady and with a further increase in blowing pressure the power absorbed decreased and the now steady shock wave moved upstream. This condition, known as 'over-blowing', was used to obtain steady near-normal shock waves in the observation section

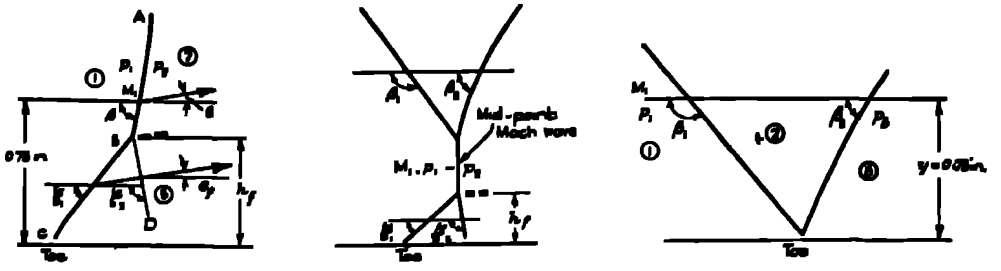


FIGURE 3 Sketches of shock wave systems (*a*) near-normal shock wave with bifurcated foot, wave strength $\chi = p_2/p_1$, (*b*) oblique incident and reflected wave system with bifurcated Mach shock wave, wave strength $\chi = p_2/p_1$; (*c*) oblique incident and reflected wave system without bifurcated Mach shock wave, wave strength $\chi = p_2/p_1$.

A 4° steel wedge fixed to the roof with the sharp edge facing upstream was used to obtain an oblique incident and reflected wave system in the observation section The effective angle of the wedge, and consequently the strength of the bow wave, was varied by tilting the wedge about its leading edge, which was always kept in contact with the roof The rear of the wedge was faired with plasticine after the incidence was set

4 2 Two methods, direct shadow and Töpler striation, were used for visualization and photography of flow phenomena. The direct shadow method was used to give a sharp well-defined image of a shock wave, needed for measurements of its shape The Töpler striation method gives a less clearly defined picture of shock-wave shape but a more detailed picture of the flow, particularly in the boundary layer The direct shadow was formed by collimating the light from a point source into a parallel beam, and projecting this beam normally through the observation section on to a screen or photographic plate

A full account of Topler's striation method is given in a paper by Schardin (1934) The optical system used in the present work is shown in figure 4 Details of the boundary-layer flow were obtained by setting the straight-edged diaphragms *C* and *F* parallel to the longitudinal axis of the tunnel. For visual observation, a high-pressure mercury arc was used as a source. For photography, a spark from an

induction coil, passed between magnesium electrodes in a glass capillary tube, was used After several sparks had been discharged the inside surface of the glass wall of the tube between the electrodes became frosted, and the tube itself became a light source giving a straight, brilliant uniform beam directed along the axis of the system A 125 cm Goerz Dogmar $f/7$ lens was used in both systems

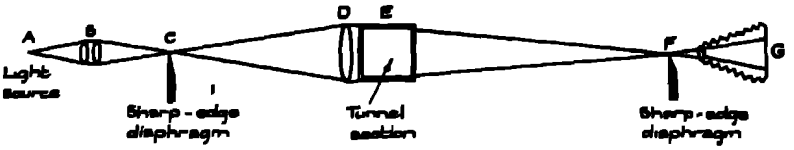


FIGURE 4 Optical arrangement for Töpler striation method.

Photographs of near-normal waves of different strength are given in figure 5*a–i*, plate 1, and in figure 6*j–l*, plate 2 Photographs of oblique incident and reflected wave systems are given in figure 6*m–q* For description of plates see p 20

4 3. The strength, χ, of a near-normal shock wave with or without a bifurcated foot, is specified by the value of p_2/p_1 at $y = 0\ 75$ in., where p_1 and p_2 are the static pressures just in front of and behind the main shock wave, see figure 3*a* that of an oblique incident and reflected wave system without a Mach shock wave by the value of p_3/p_1 at $y = 0{\cdot}75$ in , where p_1 is the pressure just in front of the incident wave and p_3 is the pressure just behind the reflected wave, figure 3*c* and that of an oblique incident and reflected wave system with a Mach shock wave and bifurcated foot by the value of p_2/p_1 taken at the mid-point of the Mach shock wave, figure 3*b* The values of χ were obtained by calculation for values of p_1 and H_1 measured in front of the shock waves and for measurements of shock-wave angles obtained from the photographs.

4 4 The sequence given by the photographs *d*, *c*, *e–i* of figure 5, plate 1, shows how the shape of a near-normal shock wave near the surface changes with the wave strength χ For a high value of χ a wave bifurcates as it approaches the surface whereas for a low value of χ it does not. The bifurcation is such that the front limb is inclined forward to the surface, whilst the back limb is shorter and inclined backward. The photographs *c–g*, $\chi \geqslant 1\ 87$, show that the height, h_f, of the bifurcated foot, i e. the height of the point of bifurcation above the surface, diminishes with a decrease in χ. Photograph *h*, taken for $\chi = 1{\cdot}84$, shows a wave about to bifurcate, and photograph *i*, for $\chi = 1{\cdot}52$, a wave that does not bifurcate. Photographs of near-normal shock waves are also given in figure 6*j*, *k*, plate 2. The first, for $\chi = 1\ 080$, the lowest value in the experiments, shows a wave which has not bifurcated, whilst the second for $\chi = 2{\cdot}33$, the highest value, shows a wave with a well-formed bifurcated foot All the photographs show, as would be expected, that the shock waves do not extend right down to the surface They also show that the layer of air passing between the toe of a shock wave and the surface thickens beyond the toe and that

the rate of thickening increases with the value of χ. This thickening is shown in more detail in the Töpler striation photographs, figures 5*a*, *b* and 6*l*. It will be noticed that the weak waves without a bifurcated foot, figures 5*h*, *i* and 6*j*, also show some thickening of the surface layer behind the toe. The toe of a wave is taken to be the lower end of the front limb, when the wave has a bifurcated foot, and the lower end of the wave, when it has not.

4·5 Direct shadow photographs of oblique incident and reflected wave systems are given in figure 6*m*, *n*, *q*. Photograph *m* shows that an oblique incident and reflected wave system has a Mach shock wave with a bifurcated foot when $\chi = 2{\cdot}25$, photograph *n* that a system has a short Mach shock wave without a bifurcated foot when $\chi = 1{\cdot}76$, and photograph *q* that reflexion occurs with a short normal wave at the foot when $\chi = 1{\cdot}39$. Values of p_1/p_2 for the incident waves of these systems were calculated for measured values of β_1 and M_1. For these values of p_1/p_2, 0·789, 0·774 and 0·811, respectively, the theoretical limiting values of β_1 for regular reflexion are 127·8, 128·8 and 126·5°, respectively. The values of β_1 obtained from the photographs are 125·1, 129·2 and 131·8°, respectively. For inviscid flow (no boundary layer), a Mach shock wave would therefore be expected for the first system but not for the other two systems. Photograph *m* and the corresponding Töpler striation photograph *o* show the rapid thickening of the layer of air which passes between the toe of the bifurcated foot and the surface. Photographs *n* and *p* show that the rate of thickening for a short Mach wave which has not bifurcated is smaller. The incident waves in all the photographs are straight, but the reflected waves are curved, especially near the point of reflexion.

5 Surface friction and pressure below shock waves

5·1 Figure 7 gives curves of static pressure p/p_A and of surface friction ratio $\tau_0/(\tau_0)_d$, where $(\tau_0)_d$ is a datum value of τ_0 measured 0·2 in. in front of the toe position, obtained from measurements taken on the floor below the near-normal shock waves shown in figures 5*d* and 6*j*–*l*, plates 1 and 2, and also for the oblique incident and reflected wave systems shown in figure 6*m*–*q*. The position of the toe of each system is marked on the curves by a vertical arrow. Figure 9*b* gives values of $\tau_0/(\tau_0)_d$ obtained, for the near-normal shock wave shown in figure 5*c*, from readings taken for the surface tube set at four heights, 0·009, 0·011, 0·012 and 0·015 in., above the surface. The readings for the different heights are seen to be reasonably consistent and accordingly the readings for the other systems were taken for the tube at one height, 0·011 in., only.

5·2. The curves of p/p_A show that the static pressure on the floor immediately behind a toe rises with an increase in the distance beyond the toe and that the pressure gradient there is quite steep for the higher values of χ. Further beyond a toe, the gradient is much more gradual and tends to be linear. The curves of $\tau_0/(\tau_0)_d$

show that the intensity of surface friction on the floor immediately behind a toe falls with the distance beyond the toe The minimum value of $\tau_0/(\tau_0)_d$ ultimately reached depends on the value of χ and is zero for the higher values of χ Table 1

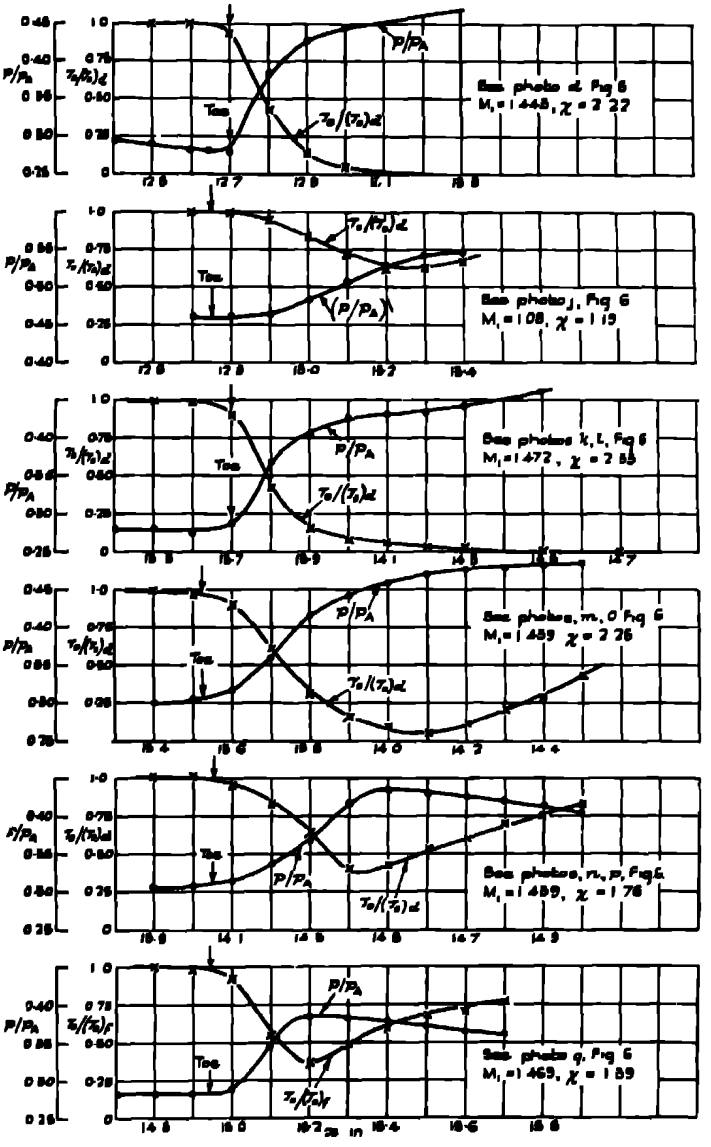


FIGURE 7. Curves of normal pressure and friction on floor below shock waves

gives minimum values of $\tau_0/(\tau_0)_d$ and also values of Δ_p/p_Δ, where Δ_p is the rise in floor pressure from the toe position to the point where $(\tau_0/\tau_0)_d$ has its minimum value, obtained from the curves of figures 7 and 9*b*.

Table 1. Shock-wave data

Type A, near-normal shock wave with a bifurcated foot, see figure 3*a*
Type B, near-normal shock wave without a bifurcated foot
Type C, oblique incident and reflected wave system with Mach shock wave and bifurcated foot, see figure 3*b*
Type D, oblique incident and reflected wave system without a bifurcated foot, see figure 3*c*

figure	photo-graph	type	tunnel bulge	x in (toe)	p_1/p_A	H_1/p_A	M_1	χ calc	measurements from photograph $\xi_1°$	$\xi_2°$	$\xi_2°-\xi_1°$	h_f in.	minimum value of $\tau_0/(\tau_0)_d$	Δ_2/p_A	$\beta°$ or $\beta_1°$ at $y = 0{\cdot}75$ in.	$\beta_2°$
5	*a**	A	II					values about the same as for photograph *c*								
	*b**	B	II	11·7	0·360	0·974	1·283	1·76	—	—	—	0	not measured		96·6	—
	c	A†	II	13·2	0·285	0·953	1·407	2·11	52·5	103·0	50·5	0·55	0	0·175	78·8	—
	d	A	II	13·3	0·288	0·976	1·443	2·22	52·5	101·0	48·5	0·46	0	0·177	80·9	—
	e	A	II	12·8	0·306	0·956	1·392	2·10	55·0	102·0	47·0	0·33	not measured		82·0	—
	f	A	II	12·1	0·330	0·967	1·346	1·95	55·0	98·5	43·5	0·14	do.		95·7	—
	g	A	II	12·0	0·335	0·971	1·336	1·87	56·5	96·0	39·5	0·14	do		100·2	—
	h	B	II	11·9	0·344	0·966	1·310	1·84	—	—	—	0	do		97·4	—
	i	B	II	11·1	0·409	0·998	1·206	1·52	—	—	—	0	do		94·6	—
6	*j*	B	I	13·7	0·501	0·960	1·080	1·19	—	—	—	0	0·63	0·078	90·0	—
	k	A	III	13·7	0·273	0·965	1·472	2·33	53·3	110·1	56·8	0·47	0	0·175	80·9	—
	*l**	A	III					values about the same as for photograph *k*								
	m	C	II	13·5	0·286 (*m*)	0·960 (*m*)	1·439	2·25‡	43·0	100·4	57·4	0·27	0·05	0·170	125·1	63·5
	n	D	II	14·0	0·286	0·960	1·439	1·76	—	—	—	0·11	0·375	0·130	129·2	61·6
	*o**	C	II					values about the same as for photograph *m*								
	*p**	D	II					values about the same as for photograph *n*								
	q	D	III	14·9	0·270	0·945	1·469	1·39	—	—	—	0·03?	0·375	0·100	131·8	50·6

* Topler striation photograph.

† Values of h_f and $(\xi_2-\xi_1)$ from three other photographs for $M_1 = 1{\cdot}407$ and $\chi = 2{\cdot}11$ are given in figure 8*a*, *b*. Mean values are $\xi_1 = 54°$ and $\xi_2 = 101°$ p_1/p_A, H_1/p_A, χ, were obtained for $y = 0{\cdot}75$ in., except ‡ which were measured at mid-point of Mach wave.

6. Shock-wave bifurcation

6·1 Values of h_f, $(\xi_2^\circ - \xi_1^\circ)$, Δ_p/p_A, $[\tau_0/(\tau_0)_d]_{\min}$ and χ for the shock waves shown in figures 5 and 6 are given in table 1 Figure 8 gives the values of h_f, $(\xi_2^\circ - \xi_1^\circ)$, Δ_p/p_A and $[\tau_0/(\tau_0)_d]_{\min}$ in this table plotted against χ Distinctive symbols are plotted to indicate the type of wave system, whilst the small letters placed near them give references to the photographs in figures 5 and 6, plates 1 and 2

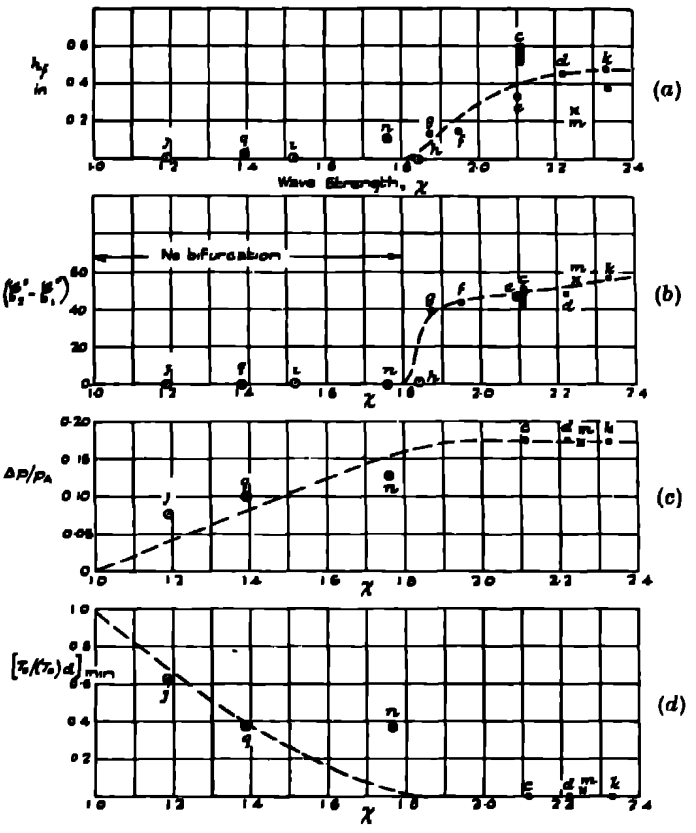


Figure 8 Dependence of properties of shock waves on wave strength. ● near-normal shock wave with bifurcated foot, ⊙ near-normal shock wave without bifurcated foot, × oblique incident and reflected wave system with Mach shock wave and bifurcated foot, ⊗ oblique incident and reflected wave system with short Mach shock wave without bifurcated foot. Small letters refer to photographs in figures 5 and 6, plates 1 and 2.

Figure 8 *a*, *b* show that the bifurcation of the foot of either a near-normal or an oblique incident and reflected wave system with a Mach shock wave does not occur until a wave strength $\chi \approx 1{\cdot}8$ is reached. Above this value of χ, both h_f and $(\xi_2^\circ - \xi_1^\circ)$ at first increase rapidly and then change slowly with χ. The four values of h_f and $(\xi_2^\circ - \xi_1^\circ)$ plotted for $\chi = 2{\cdot}11$ (c) were obtained from four different photographs and they are included to give a measure of the degree of consistency in the observations. Figure 8 *c*, *d* show that $\Delta_p/p_A = 0{\cdot}175$ and $[\tau_0/(\tau_0)_d]_{\min} = 0$ for wave

systems with a well-formed bifurcated foot, $\chi > 2{\cdot}0$. and that Δ_p/p_A increases and $[\tau_0/(\tau_0)_d]_{\min}$. decreases, but not to zero, for the waves without a bifurcated foot, with an increase in χ from 1 to 1·8.

7. Detailed explorations for a near-normal bifurcated shock wave

7·1. Information on the nature of a near-normal shock wave with a bifurcated foot was obtained from detailed explorations made for a wave system which closely resembled that shown in figure 5 *a, c*. The readings taken, together with the results of calculations made from them, are given in figure 9 *a–c* Figure 9*c* gives curves of p/p_A, H/p_A and q/a_A plotted against x, obtained from explorations made through the main shock wave at $y = 1$ and 0·75 in , and through the bifurcated foot at $y = 0{\cdot}4$ in The readings of static pressure taken immediately in front of the shock wave could not be accepted partly because the wave front itself was distorted locally by the boundary layer on the horizontal stem of the tube passing through the wave and partly because the wave had a small longitudinal movement The readings taken at a distance about 0·2 in forward of the wave front were, however, satisfactory for they were found to be in very close agreement with those taken in the supersonic stream without the wave present, see dots and crosses in figure 9*c*. The static pressure immediately in front of the wave was accordingly taken to be the same as that measured at the same position in the stream without the wave present The same procedure was used to get the stagnation pressure immediately in front of the wave No readings of static and stagnation pressures were taken immediately behind the wave because it was considered preferable to calculate them

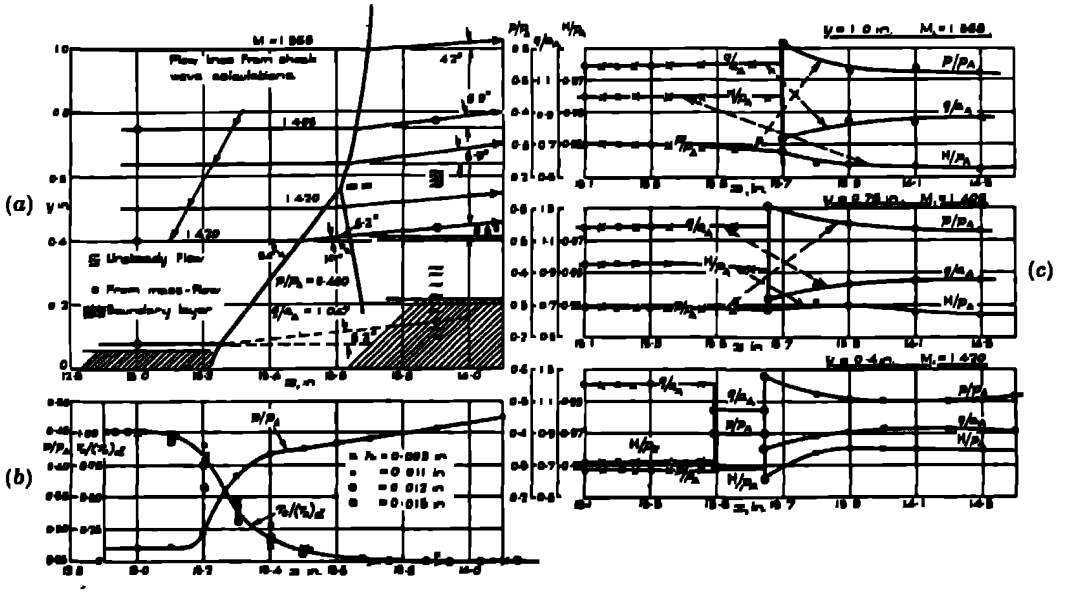


Figure 9. Properties of near-normal shock wave of type shown in figure 5*c*, plate 1. ● with shock wave in tunnel, × without shock wave, ⊙ from shock-wave calculations

by the standard shock-wave relations from readings taken in front The calculated values are shown by points in circles in figure 9*c*. they are reasonably consistent with the readings taken farther behind the wave The curves of figure 9*c* show that whereas the rise in static pressure and the fall in velocity through the bifurcated foot do not differ much from those through the main wave at $y = 0{\cdot}75$ in., the loss in stagnation pressure is about 50 % smaller

7·2 It should be mentioned that the values of the stagnation pressures H_1 estimated from readings of H_r taken in the supersonic flow in the observation section in front of a shock wave were about 4 % lower than the atmospheric pressure, p_A. There was, therefore, an increase in entropy during the flow of the air down the tunnel. This increase is due to losses of mechanical energy associated with condensation shock waves Such waves were observed in and near the tunnel throat and their nature depended on the humidity of the air outside Uncertainty arising from the effects of humidity on the observations given is, however, considered to be quite small, for observations were taken only on days of low humidity and, in general, they could be repeated with good accuracy

7·3. Figure 9*a*, *b* taken in conjunction show that the steep rise in p/p_A and the steep fall in $\tau_0/(\tau_0)_d$ along the floor occur within the two limbs of the bifurcated foot The flow lines given in figure 9*a* were obtained from shock-wave calculations It is seen that the stream passing through the main shock wave just above the point of bifurcation is deflected upwards by 6 0° and that through the bifurcated foot just below by 5·5°. The points in circles are on flow lines obtained from mass-flow calculations made for the data in figure 10 and calculated angles of deflexion. These

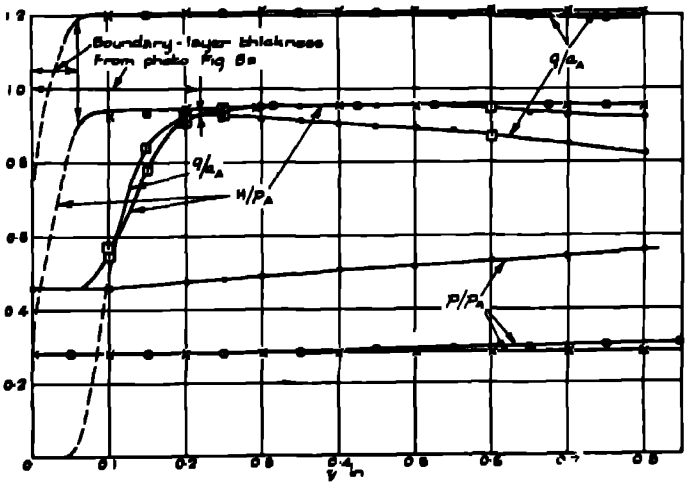


FIGURE 10 Lateral traverses made for near-normal shock wave, figure 5*a*,*c*, plate 1.

Shock-wave toe at $x = 13\ 2$ in	● steady reading.
Traverse at $x = 13\ 9$ in	□ unsteady reading.
Traverses in tunnel	× $x = 13{\cdot}9$ in
Without shock wave present	⊗ $x = 13\ 0$ in

flow lines are seen to be in close agreement with those obtained from the shock-wave calculations The shaded areas at the floor give the thickness of boundary layer before and after passing the bifurcated foot, obtained from the Töpler striation photograph figure 5*a* These thicknesses are marked on curves of H/p_A and q/a_A, obtained from pitot and static-pressure traverses, in figure 10 and it is seen that they agree reasonably well with those which would be estimated from the curves. The thickness increases from about 0·06 in in front of the wave foot to about 0·22 in. behind. The short wavy lines in figure 9*a* indicate regions behind the wave where the readings of H/p_A were unsteady, see also figure 10 These regions are behind the point of bifurcation and the lower end of the back limb.

8 Bifurcated shock-wave theory

8·1 It is convenient at this stage to consider the properties of a shock wave in an infinite uniform inviscid stream, when the wave is bifurcated at its lower end The wave system envisaged is that shown in figure 3*a*, except that there is now no floor and the front and back limbs, *BC* and *CD*, of the bifurcated foot extend to infinity Information on such a system can be obtained on the assumptions that the fluid passing through the bifurcated foot emerges with the same pressure and in the same direction as that passing through the main wave and that these two streams are separated by a vortex sheet, the velocity, density, stagnation pressure and entropy on opposite sides of the sheet being different It is obvious that the possible ranges of ξ_1 and ξ_2 for a bifurcated foot, for any assigned value of M_1, are limited, for ξ_1 cannot be smaller than $\sin^{-1}(1/M_1)$ and it cannot be greater than the angle for which the Mach number behind the front limb is unity, for, if it were, a back limb could not exist. Figure 11 gives the theoretical limits of ξ_1 within which bifurcation is possible for the range $1 < M_1 < 1{\cdot}6$. The possible range of ξ_1 increases with M_1 and

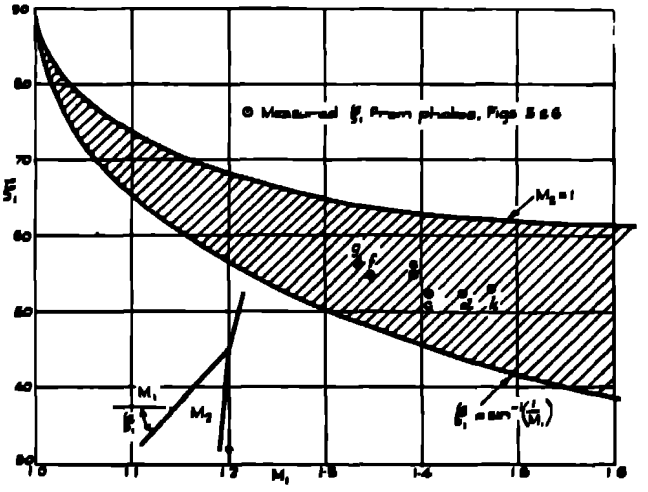


Figure 11. Theoretical limits of ξ_1 within which bifurcation is possible.

is small at low values of M_1. The values of ξ_1 measured for the near-normal systems shown in the photographs of figures 5*c–g*, plate 1, and 6*k*, plate 2, are plotted in figure 11 each value lies about midway between the theoretical limits

8·2. Figure 12 gives theoretical curves of $\xi_1°$, $\xi_2°$, $\epsilon_f°$ $(= \epsilon°)$, $\alpha_2°$, p_3/p_1 $(= p_2/p_1)$ plotted against $\beta°$ for a bifurcated shock wave in an infinite stream for which $M_1 = 1\ 45$. This value of M_1 was chosen because it is close to the values for which the well-formed bifurcated waves shown in figures 5*c*, *d*, plate 1, and 6*k*, plate 2, were obtained The possible ranges of $\beta°$, p_2/p_1 $(= p_3/p_1)$, $\xi_1°$, $\xi_2°$, $\epsilon_f°$ $(= \epsilon°)$ and $\alpha_2°$ for which bifurcation can occur are 62·7 to 98·0, 1 76 to 2 274 (max.) to 2·23, 62 7 to 43 8, 62·7 to 98·0, 10·4 to 10·6 (max) to $-5\ 75$ and 52 3 to 98 respectively The first value of each pair is that for which a back limb is just about to appear and for which the front limb is continuous with the main shock wave and the second value is that for which the front limb is just about to disappear and the back limb is continuous with the main wave The curves near the lower limit of β are shown dotted because the accuracy of the calculations made for this region is not high they are intended to indicate general trends only

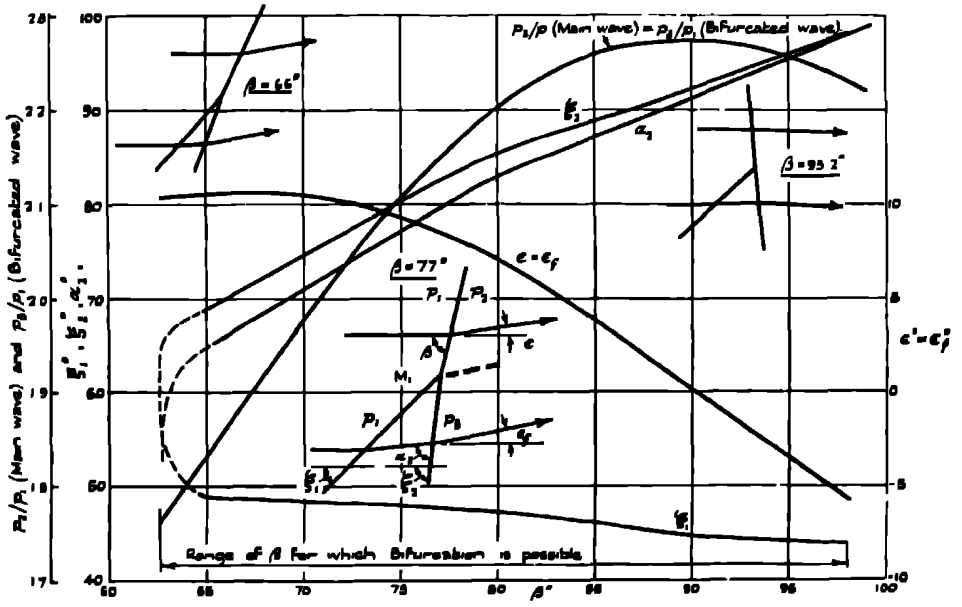


FIGURE 12. Theoretical curves for shock wave in an infinite inviscid stream, $M_1 = 1\ 45$.

8·3 Figure 13 gives the velocity and stagnation pressure behind the main shock wave and also behind the bifurcated foot for the possible range of β from 62·7 to 98° The parts of the curves for the bifurcated foot near the low er limit of β are shown dotted because the accuracy of the calculations made for this region is not high The velocity and also the stagnation pressure behind a bifurcated foot are greater than those behind the main shock wave, the differences being greatest near the middle of the β range. The loss of stagnation pressure behind the bifurcated foot is there

about one-half of that behind the main wave. Since drag is closely proportional to the loss of stagnation pressure, the drag of the bifurcated foot is also about one-half of that of the main wave.

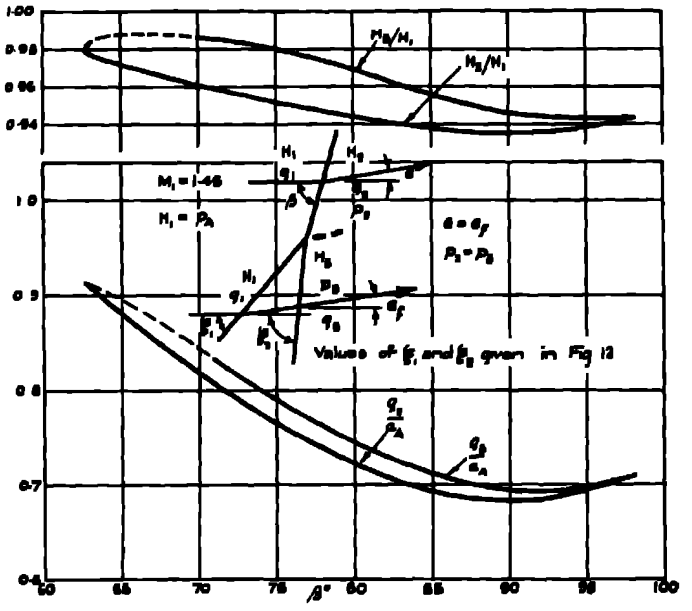


FIGURE 13. Comparison of velocity and stagnation pressure behind main wave and bifurcated foot.

9. COMPARISON OF THEORY AND EXPERIMENT. EFFECT OF BOUNDARY LAYER

9·1. The effect of the boundary layer on bifurcation phenomena at the foot of a near-normal shock wave can be illustrated by a comparison made between properties measured in the tunnel with those calculated for a main wave having the same strength in an infinite stream at the same Mach number Such a comparison, based on data given earlier in the paper, is made in table 2.

TABLE 2

		$\beta_1°$ (a)	M_1	$\xi_1°$	$\xi_2°$	$(\xi_2-\xi_1)°$	$\alpha_2°$	$\epsilon_1°$	$\epsilon_f°$
bifurcated near-normal waves in tunnel	figure 5c, plate 1	77	1·407	52 5	103	50·5	—	—	—
	figure 5d, plate 1	77	1·443	52·5	101	48·5	—	—	—
	figure 6k, plate 2	77	1·472	53 3	110	56·7	—	—	—
	means for the above systems	77	1·44	53	105	52	94·8 (b)	6·2 (b)	5·5 (b)
theory. Infinite stream	from curves of figure 12	77	1·45	47·6	82·6	35	79·7	2·9	8·5

(a) Measured just above the point of bifurcation.

(b) Calculated for the system shown in figure 9a, which closely resembles the mean of the wind-tunnel systems.

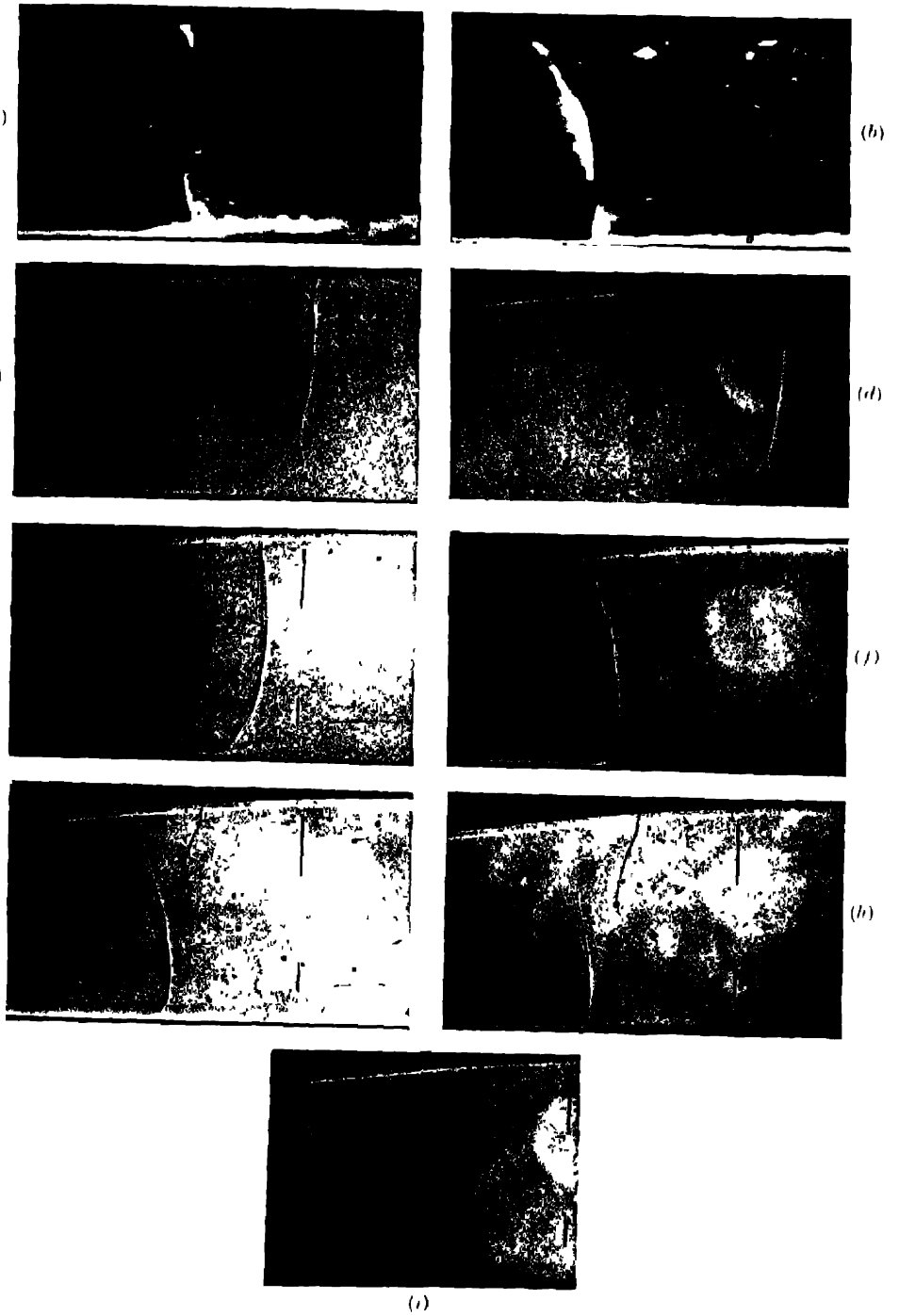


FIGURE 5

(*for description of figure see p.* 20)

(*facing p.* 16)

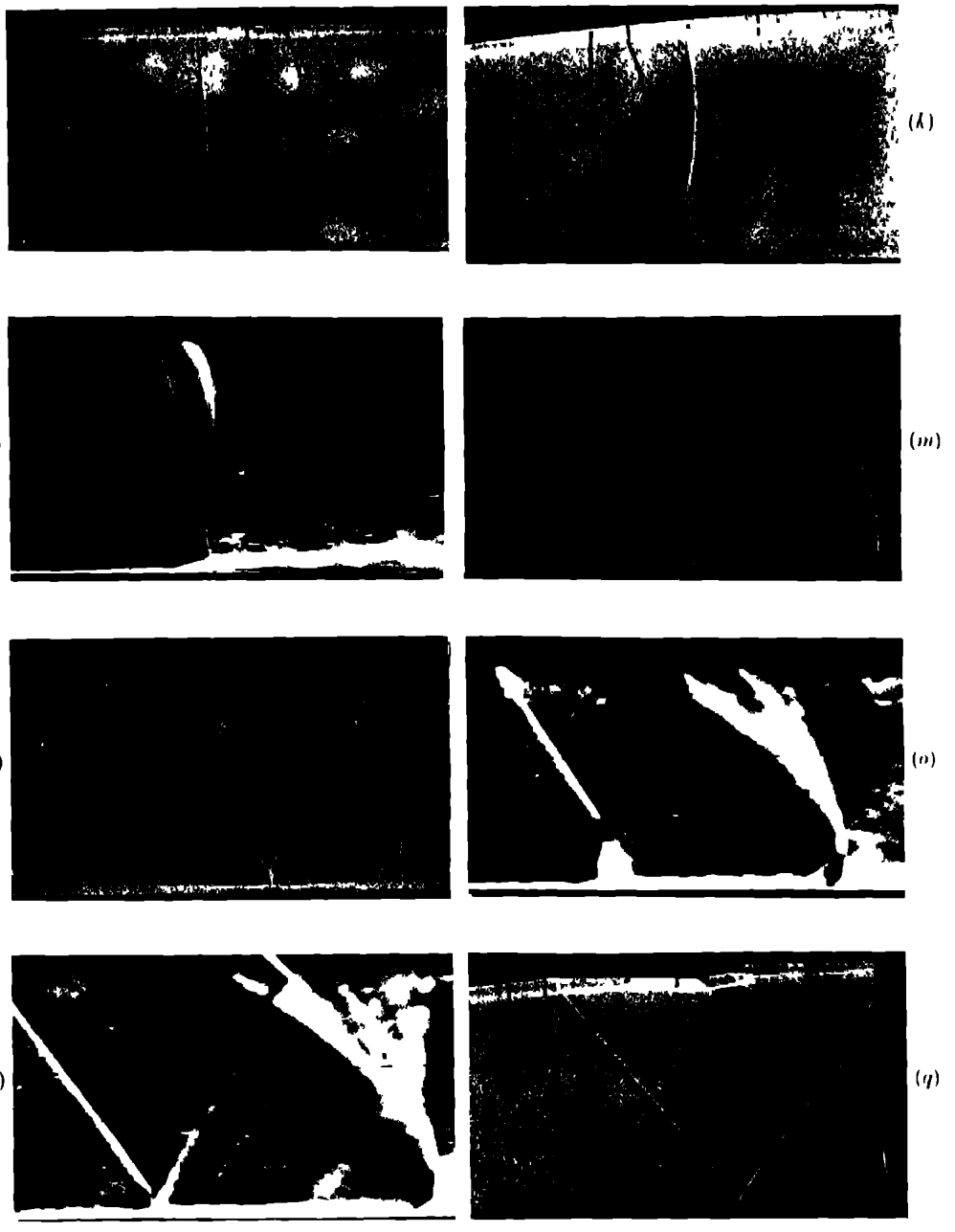


FIGURE 6

(*for description of figure see p* 20)

The presence of the boundary layer on the floor is seen to increase ξ_1, ξ_2, $\xi_2 - \xi_1$, α_2, the angle which the back limb makes with the local stream direction in front and ϵ_i, the angle of refraction through the front limb, but to decrease ϵ_f, the angle of refraction through the bifurcated foot. It is the thickening of the boundary layer behind the front limb which causes the increase in ϵ_i and so the increase in ξ_1. The angle $(\xi_2 - \xi_1)$ between the front and back limbs is then dependent on the manner in which the layer thickens between the limbs

9·2. The question now arises why the weaker near-normal waves shown in figures 5*h*, *i*, plate 1, and 6*j*, plate 2, have not bifurcated. The values of M_1 for these waves are 1·310, 1·206 and 1·080, and the possible ranges of ξ_1 given by the curves of figure 11 for an infinite inviscid stream are 50 to 65, 56 to 68 and 68 to 75° respectively. The effect of the boundary layer on the floor is to increase ξ_1 above the theoretical value for an infinite stream, so that the values of ξ_1 would have been high if bifurcation had occurred. The foot of the wave in figure 5*h* does, in fact, give a faint indication of bifurcation with the front limb inclined at a high value of ξ_1, and whilst the foot of the wave in figure 5*i* has not bifurcated into two limbs it is not only inclined at a steep angle to the surface but appears to be split into several closely spaced limbs, of which one is much stronger than the other. The shock waves of figure 5*h*, *i* were situated in the expanding flow just beyond the tunnel throat, well upstream of the positions taken by the well-formed bifurcated waves *c*, *d* and *k* The shock wave of figure 6*j* was situated farther downstream than the shock waves, figure 5*h*, *i*, and near the parallel part of the tunnel This shock wave is normal to the surface, and it is not surprising that its foot has not bifurcated because the value of M_1, 1·08, is so low. Further, the boundary layer in front has been thickened by an inclined wave upstream (see figure 6*j*, plate 2) and this thickening has probably eased the conditions of flow near the foot of the normal wave behind. Curves in figure 7 show that behind the foot of this shock wave there is a gradual rise in surface pressure and also a gradual fall in surface friction. The minimum value of the intensity of surface friction reached is about 60 % of that in front of the shock wave.

10 Conclusion

10·1 The principal results obtained from the work are given in the summary on p. 1. No attempt has been made to observe the effect of a systematic variation of turbulent boundary-layer thickness on shock-wave phenomena near a surface. It is likely, however, that bifurcation at the foot of a shock wave nearly normal to a plane surface is likely to occur at higher Mach numbers than those covered in the experiments and for thinner turbulent boundary layers. The rise of pressure through such a shock wave is high and distortion of flow near the foot appears to be inevitable since the wave can never extend right down to the surface By bifurcation at the foot and a correlated orientation of the main shock wave it is possible for a large part of the pressure rise associated with the flow through the main shock wave to occur behind the front limb of the foot and for the distortion of flow behind to be

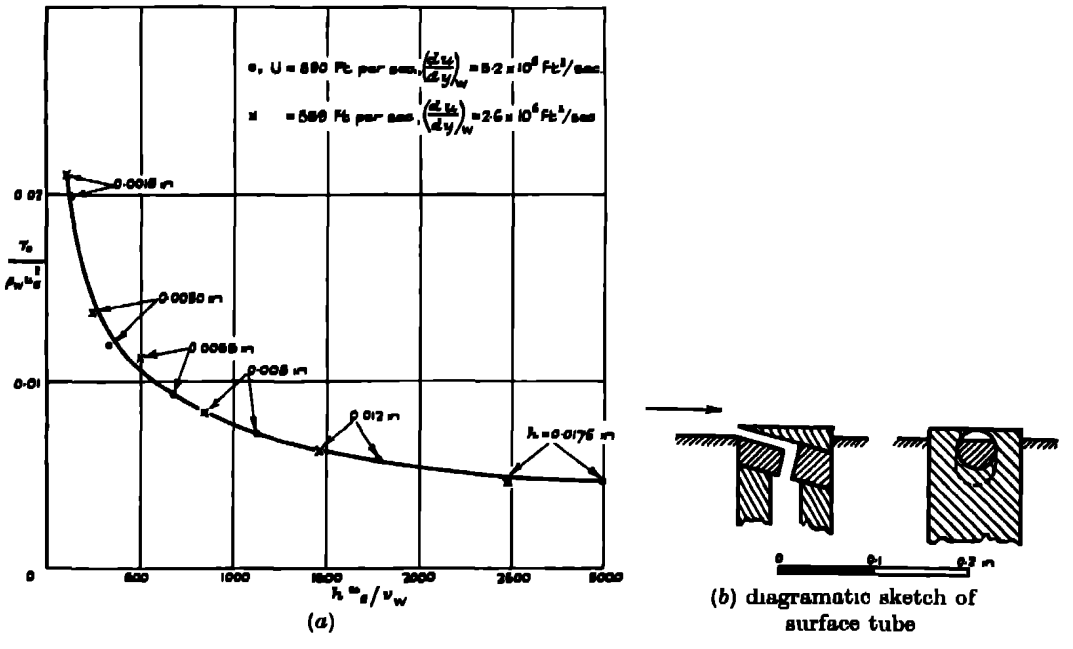
eased by the back limb with a smaller loss of stagnation pressure, and so a smaller drag, than that associated with the flow through the main shock wave. The friction on the floor in the space between the front and back limbs falls rapidly to zero so that the energy losses in the boundary layer arising from the work done against the surface friction is small. The behaviour of the boundary layer well beyond a bifurcated foot has not been considered in the present work, and no provision was made to allow this to be done.

10·2 The work described above was carried out in the Aerodynamics Division of the National Physical Laboratory, and it is published on the recommendation of the Aeronautical Research Council and by permission of the Director of the Laboratory

Appendix

Calibration of surface tube

1 The surface tube consisted of a 0·1 in circular rod with a semicircular hole, radius 0·02 in , inclined downwards to meet a 0 032 in axial hole, see figure 14*b*. The upper end of the rod was ground square to the axis until the top of the mouth had a razor edge The tube was mounted at a pre-selected height, always smaller than 0·02 in. The shape of the mouth of the tube when in position was therefore segmental The height, h, of the tube was measured with a micrometer carried on a small bridge-piece spanning the tube and resting on the tunnel floor The stagnation



(*a*) (*b*) diagramatic sketch of surface tube

Figure 14. Surface-tube calibration curve.

pressure at the tube mouth was measured against the static pressure at a small hole in the floor, situated in the same cross-section of the tunnel.

2. The value of τ_0/RU^2 at the floor was determined from a solution of the momentum equation for compressible turbulent flow along a flat plate, for distributions of U and R measured in the tunnel forward of the tube This equation is

$$\theta' + \theta\left[\frac{R'}{R} + \left(\frac{\delta_1}{\theta} + 2\right)\frac{U'}{U}\right] = \frac{\tau_0}{RU^2},$$

where the dash denotes differentiation with respect to x The relation for τ_0/RU^2 was taken to be $\tau_0/RU^2 = 0{\cdot}00653(U\theta R/\mu_w)^{-1/6}$ The values of δ_1/θ taken were derived on the assumption that the velocity profile was of the form

$$\frac{u}{U} = 1 + \frac{1}{K}\frac{U_\tau}{U}\log_e\left(\frac{y}{\delta}\right),$$

the estimated value of $\frac{1}{K}\frac{U_\tau}{U}$ being 0 096 The approximate value of δ_1/θ is then given by $1{\cdot}24(1+0{\cdot}36M^2)$, and that of δ_1/δ by $0\,096(1+0\,30M^2)$ The results of the boundary-layer calculations are given below·

U f p s	$M = U/a$	R/ρ_A	θ in	δ_1/θ	θ/δ	δ in.	τ_0/RU^2
890	0 855	0 714	0 0223	1 57	0 075	0 30	0 00148
559	0 528	0 882	0 0262	1 37	0·077	0 34	0 00150

3 Figure 14*a* gives values of $\tau_0/\rho_w u_e^2$ obtained from the surface-tube calibrations plotted against hu_e/ν_w, where suffix w denotes values at the wall and u_e is the effective velocity at the tube mouth calculated for the measured stagnation and static pressures. Results obtained for $U = 890$ f.p s. are in close agreement with those obtained for $U = 559$ f.p s

The shape of the tube mouth changes with h, so that $\tau_0/\rho_w u_e^2$ is a function not only of hu_e/ν_w but also of h/b, where b is the width of the mouth at the surface. The calibration curve given in figure 14*a* should, however, give a reasonably good indication of changes in $\tau_0/\rho_w u_e^2$, and so of $(du/dy)_w$, provided that the conditions under which the tube is used do not differ much from those of calibration This form of surface tube was used because of the ease with which it can be made with a razor edge at the top wall. The dependence of the reading on h/b as well as on hu_e/ν_w would have to be determined before the tube could be used for absolute measurement of surface friction It was estimated that the thickness of the viscous layer at the calibration position was about 3×10^{-3} in and that of the laminar sublayer about 3×10^{-4} in.

Reference

Schardin 1934 *Forschung auf dem Gebiete des Ingenieurwesens.* Forschungsheft 367.

Description of Plates 1 and 2

Plate 1

Figure 5. (*a*) to (*b*) Töpler striation photographs; (*c*) to (*i*) direct shadow photographs

(*a*) $M_1 = 1\ 407$, $\chi = 2\ 11$ (*d*) $M_1 = 1\ 443$, $\chi = 2\ 22$ (*g*) $M_1 = 1\ 336$, $\chi = 1\ 87$.
(*b*) $M_1 = 1{\cdot}283$, $\chi = 1\ 76$ (*e*) $M_1 = 1\ 392$, $\chi = 2{\cdot}10$ (*h*) $M_1 = 1\ 310$, $\chi = 1{\cdot}84$
(*c*) $M_1 = 1{\cdot}407$, $\chi = 2\ 11$ (*f*) $M_1 = 1\ 346$, $\chi = 1{\cdot}95$. (*i*) $M_1 = 1\ 206$, $\chi = 1\ 52$

Plate 2

Figure 6 (*l*), (*o*), (*p*) Töpler striation photographs; (*j*), (*k*), (*m*), (*n*), (*q*) direct shadow photographs

(*j*) $M_1 = 1\ 08$, $\chi = 1\ 19$ (*m*) $M_1 = 1\ 439$, $\chi = 2\ 25$. (*p*) $M_1 = 1\ 439$, $\chi = 1\ 76$.
(*k*) $M_1 = 1\ 472$, $\chi = 2{\cdot}33$ (*n*) $M_1 = 1\ 439$, $\chi = 1\ 76$ (*q*) $M_1 = 1\ 469$, $\chi = 1{\cdot}39$.
(*l*) $M_1 = 1\ 472$, $\chi = 2{\cdot}33$ (*o*) $M_1 = 1\ 439$, $\chi = 2\ 25$

The assignment of the slow-neutron-produced activities of thallium and the dual disintegration of radium E

By E Broda and N Feather, F.R.S.

(*Received* 5 *December* 1946)

Recoil experiments with a strong source of RaE (21 mC initial activity) electrolytically deposited on platinum show that an activity of half-value period 4·2 min is obtained. This observation indicates that Ra E undergoes dual disintegration with an $\alpha : \beta$ branching ratio of the order of 5×10^{-7}. Rough absorption measurements favour the identification of the activity with the 4·2-min activity produced in the (n, γ) and (d, p) reactions with thallium, and therefore require assignment of the latter to ^{206}Tl, contrary to earlier suggestions. Consideration of the energies involved shows that the postulated α : β branching ratio has roughly the value which would be predicted on the basis of accepted regularities.

Irradiation of thallium by slow neutrons, as well as its transformation in the (d, p) reaction (Fajans & Voigt 1941), leads to the production of β-activities of half-value periods 4·2 min and 3·5 years. On the basis of the observation of Heyn (1937), that the 4·2-min. activity is also obtained as the result of fast-neutron irradiation, it has generally been assumed (see, for example, Seaborg 1944) that the 4·2-min activity is to be attributed to ^{204}Tl and the activity of 3·5 years half-value period to ^{206}Tl. The arbitrary nature of this assignment has been pointed out by Fajans & Voigt (1941), and it is the object of the present paper to indicate how the matter may be further investigated from a different starting-point and to describe some experiments, in which it was possible for ^{206}Tl but not ^{204}Tl to be produced, which show

conclusively that the currently accepted assignment is incorrect. In these experiments the 4·2-min. activity was in fact obtained, but not the other, thus as a result of them it must be concluded that the short-period activity definitely belongs to the heavier, and the long-period activity presumably to the lighter, of the two thallium isotopes in question Hoyn's observation must then be explained either as an early example of the detection of fast-neutron capture, or on the basis of his failure completely to exclude slow neutrons in his experimental arrangement.

Briefly, the new experimental approach is by way of a close reinvestigation of the natural radioactivity of Ra E ($^{210}_{83}$) Obviously, if α β branching occurs with Ra E, as with the other known β-active isotopes of bismuth, then $^{206}_{81}$Tl will be formed in the rare mode of disintegration Energetically, α β branching is possible so long as the energy liberated in the β-disintegration of ^{206}Tl is less than 6 57 MeV, the sum of the disintegration energies of Ra E and Ra F, practically, its observation is likely only if the ^{206}Tl β-disintegration energy lies within fairly narrow limits We shall inquire, therefore, what the branching ratio would be, first on the assumption that the 3·5-year body is to be identified as ^{206}Tl and then on the assumption that the 4·2-min. activity is to be assigned to this species The accepted values for the β-disintegration energies are 0 87 MeV in the former case (Fajans & Voigt 1941) and 1 70 ± 0 05 MeV (Fajans & Voigt 1940, 1941, Krishnan & Nahum 1940) in the latter, thus the energy released in the hypothetical α-disintegration of Ra E is 5·70 MeV according to the first and 4·87 ± 0 05 MeV according to the second assumption. Figure 1 provides the basis for an estimate of partial disintegration constants once

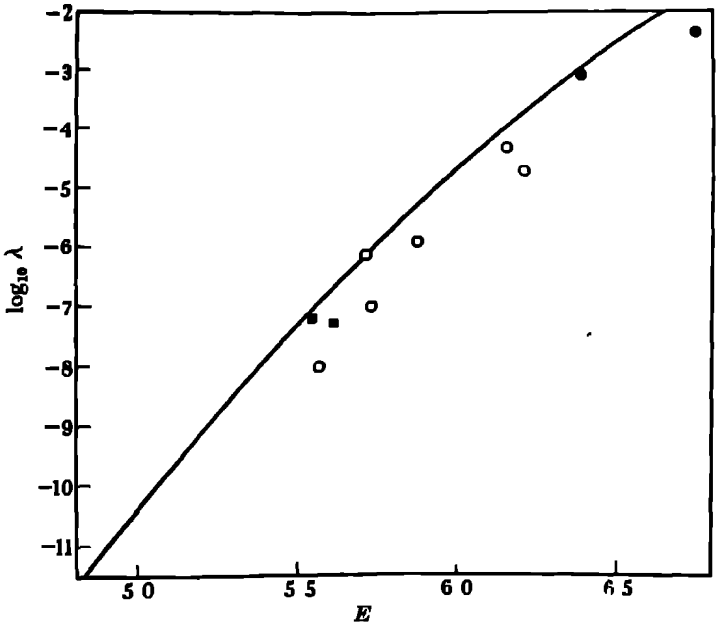


FIGURE 1. ● Ac C, ○ Th C, ■ Ra C

these energies are given. In this figure total disintegration energies and partial disintegration constants are plotted (the latter logarithmically) for all the α-particle groups belonging to the known modes of disintegration of Ac C ($^{211}_{83}$), Th C ($^{212}_{83}$) and Ra C ($^{214}_{83}$). The full curve bounding the points on the side of λ greatest (drawn and extrapolated with reference to corresponding curves for α-bodies having other values of Z (Feather 1946))* is assumed to represent the relation between $\log_{10}\lambda$ and E† for disintegrations in which there is no change of spin—and the distribution of points below the curve indicates the extent to which spin changes are likely to result in smaller values of λ (by a factor of roughly 10 in the extreme cases illustrated) than would otherwise apply From this figure, using the energies already quoted, table 1 has been prepared From the values of the α β branching ratios given in the table it is evident at once that the assignment of the 3 5-year activity to ^{206}Tl runs entirely contrary to the crudely determined fact that there is no very obvious α-component in the total radiation from Ra E—at least so long as our general views on the probabilities of radioactive change remain unchallenged But it is also clear from the table that, if the alternative assignment were to be adopted, and the β-particles of 1 70 MeV limiting energy were to be attributed to the disintegration ^{206}Tl (4·2 m) $\xrightarrow{\beta}$ ^{206}Pb, then there would be nothing very surprising in the fact that the $\alpha:\beta$ branching of Ra E has not hitherto been discovered—or in the fact that the feeble higher-energy component due to the presence of ^{206}Tl in equilibrium in Ra E sources has so far remained undetected The experiments of Bastings (1924) on the decay of strong sources, as measured by the ionization produced by the β-particles of highest energy only, would not have been sensitive enough to detect any initially anomalous decay, even if his sources had been rapidly prepared and examined, and the most careful investigations on the upper limit of the β-spectrum by the magnetic method (such as those of Alichanian, Alichanov & Dželepov (1938)) would fail, on account of background effects due to scattered β-particles, to indicate the presence of the 1·70 MeV component if its intensity were less than 10^{-3} with respect to the main spectrum having its end-point at 1·17 MeV. It would, in fact, appear that, of the experiments reported up to date, those of Gray & Henderson (1936) by the simple absorption method set the lowest limit to the possible admixture of higher-energy particles with the main radiation from Ra E. These authors claim to have shown that not more than one β-particle in 2×10^6 has sufficient energy to penetrate 0·400 g./cm.^{2} aluminium plus 0·212 g./cm ^{2} lead. Accepting this figure, it can readily be granted that the degree of admixture of a 1 70 MeV component with the main radiation cannot be greater than 1 in 10^4. Summing up the position, then, one would say that the negative evidence from previous experiments is overwhelmingly in favour of the attribution of the 4·2-min thallium activity to ^{206}Tl and the 3 5-year activity to ^{204}Tl, rather than the reverse, which, as already stated, is the currently

* From back numbers of periodicals recently received, it appears that Berthelot (1942) pointed out the regularities which are evident in Geiger-Nuttall diagrams plotted for individual values of Z before this point was stressed by one of the present writers.

† λ in sec $^{-1}$, and E in MeV.

accepted attribution Further, it may be repeated in anticipation, the new experiments now to be described provide positive evidence for the correctness of this conclusion.

TABLE 1

assumed half-value period of ^{206}Tl	energy of α-disintegration of ^{210}Ra E	α β branching ratio for Ra E
3·5 years	5·70 MeV	37:63 to 3·7·96 3
4 2 min.	4·87 ± 0·05 MeV	9×10^{-6} to 2×10^{-7}

Our first experiments were designed to test whether the 3·5-year body is in fact produced from Ra E, unprejudiced by the consideration that a positive result would be extremely unlikely for the reasons already set forth We had at our disposal a solution containing about 10 mC Ra D which had not been treated chemically for at least 8 years. This solution was evaporated to dryness, and the residue, which contained a few milligrams of inactive lead, was dissolved in HCl and small quantities of thallous and didymium salts were added as carriers. Didymium hydroxide was precipitated with ammonia and this precipitate carried down Ra D, Ra E and polonium. The process of addition of didymium and precipitation with ammonia was repeated until the filtrate was considered to be sufficiently free from Ra E for the purposes of the test Ammonium iodide was then added to precipitate the thallium The mixture was filtered and the precipitate of thallous iodide was tested for activity using a G.M. counter having a mica window of about 3·5 mg /cm ^{2} thickness covering one end A small activity was observed initially, but when the precipitate was dissolved and reprecipitated, after traces of lead and bismuth salts had been added to the solution, the residual counting rate due to the thallium was reduced to 0 8 ± 0·4 min.$^{-1}$ under the geometrical conditions obtaining in the experiment. In order to make sure that the negative result of this test was not due to the fact that the active thallium had been preferentially adsorbed on the glass of the vessel in which the Ra D solution had been standing, an acid solution of thallium salt was brought to the boil and allowed to stand in this vessel overnight and a similar test was made on the iodide precipitate obtained from this solution A similarly negative result was again found Taking count of the possibility of incomplete radioactive equilibrium in the parent solution, and of losses in the chemical processes involved, these results set an upper limit to the α β branching ratio corresponding to the formation of the 3·5-year thallium from Ra E at a value certainly not greater than 10^{-8}. 1. This conclusion having been reached, attention was concentrated on testing for the formation of the 4 2-min active thallium, as described below

For this second stage of the investigation, chemical methods of separation were tried first A procedure was evolved which allowed of the separation of thallium with a reasonable yield (25 to 50 %), even from a very large excess of bismuth (with Ra E), and the production of a sample in a form ready for counting, in 7 or 8 min Before the method could be applied the polonium content of the main solution had first to be reduced as far as possible. To this end the Ra (D + E + F) was taken up in

N/10 HCl and three copper foils were rotated for successive periods of 1 hr. each in the solution The last traces of polonium were removed by adding about 1 g Bi_2O_3 dissolved in HCl, some crystals of telluric acid, and finally a solution containing about 200 mg $SnCl_2$. The polonium and tellurium, which precipitate quantitatively in the cold, were removed by filtration The filtrate was evaporated to dryness, dissolved in 1 1 HCl, and extracted with ether repeatedly until no further solid was recovered. The solution, now ready for the thallium separation, contained, at the time of the first experiment, about 7 mC Ra E and 0·1 mC polonium. The method of separation was as follows To the solution of bismuth (and Ra E) chloride, about 12 mg $TlNO_3$ was added, together with excess bromine to oxidize the thallium, and the whole was heated on a water bath. After the $TlNO_3$ had dissolved and the colour of the bromine had more or less disappeared, the solution was chilled with ice and stirred quickly with 2 c.c of ether for about 10 sec The end of the stirring was taken as the effective time of separation of the radiothallium. The ether, containing the extracted thallium chloride was then pipetted off, washed four times with 1 1 HCl in a separating funnel, and evaporated drop by drop on a warm glass dish The solid residue so obtained was transferred to the counter and tested After the method had been perfected the initial activity of the thallium residue was generally of the order of 10 to 15 counts per minute (counter efficiency ~ 3 %) With such a small counting rate it was naturally difficult to investigate a possible short-period decay, but even to attribute the whole observed effect to ^{206}Tl is already to set quite a low upper limit to the branching ratio concerned (not greater than 10^{-6} 1, assuming 25 % efficiency for the chemical processes involved) As to the decay of the effect, the conclusion from a large number of extractions was definitely that a decrease in counting rate took place during the first few minutes' observation, but it was evident from control experiments that part of this decrease at least was associated with changes in temperature (the sample being presented to the counter on a dish which had just been heated) and the final result was that the experiments, though suggestive, were at best inconclusive

The attempt was then made to separate an active thallium from Ra E by the method of recoil To this end it was necessary to prepare a strong source of Ra E of the minimum thickness, and containing as little polonium as possible in order to reduce aggregate recoil effects to tolerable proportions A solution of Ra (D + E + F) of about 50 mC strength was freed from most of the polonium by the rotation of six copper foils. Thereafter two nickel foils were rotated in the solution for 45 min each, the acidity being kept at N/10 HCl and the temperature that of the water-bath. The nickel foils were dissolved in nitric acid, about 200 mg. didymium nitrate was added, and didymium and Ra E were precipitated as hydroxides with ammonia. The precipitate was then washed with ammonia, to remove all traces of nickel, and was dissolved in trichloracetic acid. This solution was evaporated to dryness and the residue again taken up in trichloracetic acid (2 %) and electrolyzed for 12 hr., using platinum electrodes, each of area about 7 $cm.^2$, at 1 9 V and 1 mA. About 21 mC of Ra E was found on the cathode after electrolysis. The weight of the deposit was

certainly not more than 0·1 mg and may have been considerably less, the surface of the platinum showing no discoloration

Recoil experiments were made with this source over a period of 8 days. Throughout these experiments the source remained in position in the recoil vessel, and its decay was followed by removing the ebonite plug from the vessel, replacing it by a brass plug bored with a 0·25 in hole and carrying an absorber of aluminium of 125 mg /cm ^{2} thickness, and bringing the whole vessel (with source) into a standard position below a β-ray electroscope The activity of the source, measured in this way, was found to decrease accurately according to an exponential law with a half-value period of 5·0 days Exposures were made using a brass button of 2 0 cm. diameter as collector, the button being held at a distance of 6 mm from the platinum disk carrying the Ra E, and the collecting voltage being 450 V Both for the investigation of the possible short-period decay of the activity collected by recoil, and, when this had been established, for the study of the absorption of the radiations from the short-lived body so formed, as well as for the investigation of the variation of the amount of short-period activity collected with the age of the primary source, sets of six exposures of 12 min. each were made, with just sufficiently long intervals in between for a study of the initial decay of the activity to be made. Table 2 is a summary of the results obtained with the first set of six exposures, which established the fact that a short-lived body of half-value period about 4 2 min. was in fact collected The figures quoted in the table are the total numbers of counts recorded (the 'natural' of the counter being included) during the periods of counting given in the left-hand column, for each of the six sources belonging to the set in question. That the 'residual' count (19 to 27 min from the end of the exposure) is systematically greater for sources 2, 4 and 6 than for 1, 3 and 5 is simply to be explained by the fact that the exposure button was newly cleaned only before exposures numbers 1, 3 and 5, for the alternate exposures it was put back to expose with the decayed active deposit from the previous exposure still adhering to it The totals given in the extreme right-hand column of the table were treated numerically on the assumption that a simple admixture of a constant background counting rate (counter 'natural' plus the activity due to aggregate recoil) with an exponentially decaying rate was involved. In order to determine the background rate it was assumed that the short decay period was in fact 4 2 min. On this basis the total of 1740 counts for the set of six observation periods 19 to 27 min after the end of the exposure was assumed to be made up of 81 ± 4 counts due to the short-period product which had almost decayed and 1659 ± 28 due to background Of the mean rate due to this background ($34\,6 \pm 0{\cdot}6$ min.$^{-1}$) more than two-thirds (about 24 min $^{-1}$) was known to be due to counter 'natural') The excess (about 10 min $^{-1}$) represents the activity collected by aggregate recoil—effectively during 18 min. exposure.* The nature of this activity was examined by longer exposures as will be described below.

* Taking count of the fact that the button was cleaned for re-exposure on three occasions only (vide supra).

Having determined the background rate as just described, appropriate deductions were made from the totals for the counts belonging to the eight successive sets of 2 min. periods of observation to which the entries in the main body of table 2 refer. In figure 2 these corrected totals, with probable errors, are plotted logarithmically against the mean times of observation, referred to the end of exposure as $t = 0$.

TABLE 2

time from end of 12 min exposure (min)	source						
	1	2	3	4	5	6	totals
1 – 3	183	175	175	162	148	159	1002
3½– 5½	145	141	129	129	117	122	783
5½– 7½	122	129	107	136	101	113	708
7½– 9½	100	133	94	82	83	96	588
10 –12	101	110	92	89	88	112	592
12½–14½	92	79	81	98	79	84	513
14½–16½	94	105	78	82	62	77	498
16½–18½	89	80	69	87	72	88	485
19 –27	260	338	263	297	263	319	1740

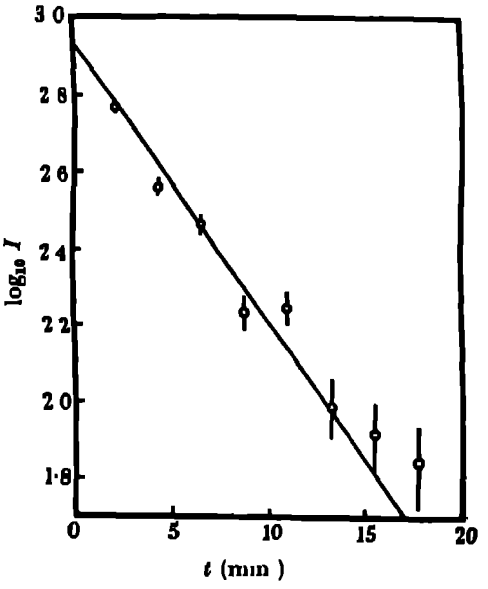


FIGURE 2

The straight line in the figure represents an activity with half-value period 4·2 min. and with most probable initial strength as deduced from the numerical analysis already carried out. The lie of this line amongst the points is sufficient proof that the whole basis of interpretation is essentially correct. A short-period activity of 4·2-min period is very obviously in question. At the mean time at which the set of six ex-

posures here dealt with was made, the calculated counting rate due to the amount of this short-lived body which was collected by recoil in 12 min. under our standard conditions was $(69{\cdot}7 \pm 3{\cdot}1)$ min.$^{-1}$ at $t = 0$

The results just analysed refer to the first set of six exposures made when the primary Ra E source was at its strongest Table 3 gives also the final result obtained from later sets, again expressed in terms of the calculated counting rate under standard conditions at $t = 0$, deduced by similar numerical analysis of the other observations. The first row of the table indicates the number of the experiment, the second row shows the mean activity of the primary source in electroscope divisions per minute, determined as already described, whilst the third row gives the corresponding calculated initial rate due to the collected activity in counts per minute. Relative values of the yield per atom of Ra E disintegrating are then given (row 4) as the quotients obtained by dividing the numbers in row 3 by the corresponding numbers in row 2 of the table It will be observed that, within the limits of error of the comparison, the yield is constant when referred to unit quantity of Ra E in the primary source. This result is clearly an unambiguous proof that the activity collected is that of an immediate daughter product of Ra E (in a very rare mode of disintegration) and that its production is in no way explicable by the initial presence or later growth of polonium in the source (production of an unidentified radioelement by α-particle-induced transformation in the air or the electrode material—or formation of ^{206}Pb in a metastable state in a rare mode of disintegration of polonium itself)

TABLE 3

experiment	1	2	3
mean activity Ra E	103 1	88 1	38·4
calculated initial activity of 4·2-min. body	69·7 ± 3 1	59 0 ± 3 9	23·3 ± 2 7
relative yield	0 676 ± 0 030	0 670 ± 0·044	0 606 ± 0 070

Absorption experiments, already briefly mentioned, provided further evidence for the identification of the short-lived body collected by recoil with $^{208}_{81}$Tl—or rather with the 4·2-min active thallium of earlier investigators In these experiments attention was concentrated on obtaining a fairly close estimate of the reduction in counting rate due to a single absorber, and for this purpose an aluminium foil of 78 mg /cm ^{2} thickness was chosen, as representing about one-tenth of the effective range of the β-particles from the short-lived thallium in question (Fajans & Voigt 1940; Krishnan & Nahum 1940). With each of a set of six sources obtained by recoil (12 min. collection) alternate periods of 2 min counting were carried out with and without absorber, during the first 18½ min life of the source (cf. table 2), and then two successive periods of 8 min., with and without absorber, for the estimation of the background effect. Numerical analysis was carried out as before The final result was that the reduction factor for the 78 mg./cm.^{2} foil in relation to the β-particles from the 4·2-min. body was found to be 0 685 ± 0·062. By comparison with the absorption

of the primary β-particles of uranium X_2, this is just about what would be expected for the β-radiation from an 'allowed' transition using an absorber of thickness roughly one-tenth of the effective range of the particles.

It has been mentioned that, in addition to the 4·2-min activity, collection by the method of recoil yielded a background activity which, in our numerical analyses, has been assumed constant over the period covered by the experiments with each short-exposure source This point was checked by taking longer exposures, which showed that the amount of background activity was directly proportional to the time of exposure (for exposures made within a few hours of one another), and by decay and absorption measurements on the radiations from the stronger sources of background activity collected during exposures lasting overnight. These measurements showed that the substance responsible for the background activity had very closely the same properties as the Ra (E + F) mixture constituting the parent source at the time the exposure was made To attribute the effect to aggregate recoil, therefore, was the natural assumption Aggregate recoil in this case is certainly due almost entirely to the presence of polonium in the source, and it is in line with this assumption that it was found that the intensity of the background activity collected by overnight exposure did not decrease anything like so rapidly as the activity of the Ra E contained in the parent source. Assuming that the aggregate recoil effect as observed in our conditions of counting, and for a constant exposure time, is proportional to the product of the mean activities of Ra E and polonium in the parent source during the period of exposure, we have the result that the intensity of the effect for short exposures should decrease or increase initially with increasing age of the primary source as the number of atoms of polonium in the source initially is greater or less than $\lambda_1/(\lambda_1+\lambda_2)$ of the initial number of Ra E atoms present Here λ_1 and λ_2 are the disintegration constants for Ra E and polonium, respectively In our case the amount of background activity collected per minute exposure time decreased during the first few days' life of the parent source. On these assumptions, therefore, the initial *activity* of the polonium in the source must have been greater than $\frac{1}{25}$ of that of the Ra E, that is greater than 0 75 mC. This is not entirely surprising in view of the time taken to work up the original source of 50 mC strength. At no time, however, was aggregate recoil a serious difficulty in our main experiment. The counting rate due to the background effect was never greater than 0·6 per min. for each minute of exposure—and as exposures in general were of 12 min duration it will be seen that the overall result of aggregate recoil was effectively to increase the 'natural' of the counter by about 30 %.

Returning to the question of branching ratio, since our standard exposure of 12 min. collected only 0 862 of the short-period activity which would have been collected during an 'infinite' exposure, we conclude that the saturation activity of the 4·2-min. body in the parent source at the time of experiment 1 of table 3 must have been at least $(2 \times 69{\cdot}7)/0{\cdot}862$ counts per minute in our counter arrangement (collection by the method of recoil cannot be more than 50 % efficient). We determined the efficiency of the counter in this arrangement for the primary β-particles

of Ra E, obtaining the value $(3\,4 \pm 0{\cdot}1)\,\%$ Taking the strength of the parent source at 17 5 mC at the time of the experiment in question, we obtain, therefore, as a lower limit to the branching ratio

$$\frac{2 \times 69{\cdot}7 \times 100}{0\,862 \times 3{\cdot}4 \times 60} \times \frac{1}{17\,5 \times 3\,7 \times 10^{7}} = 1{\cdot}2 \times 10^{-7}$$

Remembering that the result of the chemical experiment led to a 'safe' upper limit of 10^{-6} for this ratio, we may be satisfied that we now know the ratio within a factor of 5 and also that our strong source of Ra E prepared by electrolysis was reasonably 'clean' in relation to α-recoil, that is that the mean surface density of radioactive deposit was less than 0·01 mg /cm ^{2} of 'heavy' elements or about 0 002 mg /cm ^{2} of 'light' elements It will be noticed that the experimental value for the branching ratio is probably within the range of values given in table 1, though it is rather low in that range If it is assumed that the disintegration energy of the 4·2 min body is known to the accuracy supposed, two explanations are possible either our extrapolation of the $(\log_{10}\lambda)/E$ curve is somewhat at fault or is not strictly relevant (which is by no means excluded) or the α-disintegration of Ra E involves a change of nuclear spin. It might be remarked that there is every reason to suppose that the latter is in fact the case (the β-disintegration of Ra E is 'forbidden', that of the 4·2-min body is 'allowed', and the α-particle disintegration of polonium is unlikely to involve any spin change). On this supposition, then, the findings of experiment and the theoretical predictions are in full accord

Regarded from this point of view, our discovery of the generic relation between Ra E and the 4·2-min thallium is in effect a determination, within narrow limits, of the energy emitted in the disintegration

$$\binom{210}{83} \xrightarrow{\alpha} \binom{206}{81}.$$

This energy appears to be $4{\cdot}87 \pm 0{\cdot}05$ MeV. It is interesting to compare this energy with the α-disintegration energies of the other isotopes of bismuth ($Z = 83$) Values are listed in table 4 As the mass number decreases the energy of α-disintegration at first increases steadily and then, through $A = 210$, falls abruptly until with $^{209}_{83}$Bi it is so small that the radioactivity of natural bismuth has never been detected This variation of E with A for Z constant is reproduced exactly for $Z = 84$. Disintegration energies are known for $A = 218$, 216, 215, 214 and 212, and for this series there is once more a steady increase of E as A decreases, then, even more rapidly, E decreases again for $A = 211$ and decreases still further for $A = 210$ (polonium)*. For $Z = 85$ information is less complete, but for this element the effect is again evident; the α-disintegration energy for $A = 211$ is definitely less than for $A = 218$ and 216. Since for $Z > 83$ no stable isotope is known, the rapid decrease in α-

* This effect previously remarked by one of us (Feather 1946), and referred to as the polonium anomaly, had at an earlier date been commented on by Meyer (1936) whose paper containing these comments had escaped our notice

disintegration energy beyond a certain value of A (as A decreases in each case) can only be understood on the assumption that the lighter isotopes of these elements are unstable in relation to positron (or K-capture) disintegration—or that they undergo fission spontaneously at a relatively high rate. We already know, in fact, that the former suggestion is valid for ($^{211}_{85}$) which exhibits α: K-capture branching with a ratio of 40 60, but this does not necessarily invalidate the latter suggestion in other possible cases.

TABLE 4

mass number, A	214	212	211	210	209
energy of α-disintegration MeV	5 61	6 20	6 74	4 87	stable

Finally, in regard to nomenclature we would suggest that since ^{206}Tl has now been shown to be a member of the uranium series of elements, its designation in that series should be Ra E″—bringing it into line with the other thallium isotopes which are derivatives of the 'classical' radioelements.

Throughout the whole of the work described here we have had the constant assistance of Miss P. K. Wright, both in the chemical work and in the routine counting of 'aliquots' for calibration purposes Our best thanks are due to her for this help

The work was done at the Cavendish Laboratory, Cambridge, during the period March to June 1945, and is now published with the concurrence of the Ministry of Supply (Department of Atomic Energy)

REFERENCES

Alichanian, Alichanov & Dželepov 1938 *C R Acad Sci U R S.S* **19**, 375.
Bastings 1924 *Phil Mag.* **48**, 1075
Berthelot 1942 *J. Phys Radium*, **3**, 52.
Fajans & Voigt 1940 *Phys. Rev* **58**, 177
Fajans & Voigt 1941 *Phys Rev* **60**, 619
Feather 1946 *Proc. Roy. Soc Edinb.* **62**, 211.
Gray & Henderson 1936 *Phys Rev* **49**, 477
Heyn 1937 *Nature*, **139**, 842
Krishnan & Nahum 1940 *Proc Camb Phil Soc* **36**, 490.
Meyer 1936 *S B. Akad. Wiss. Wien*, IIa, **145**, 577
Seaborg 1944 *Rev Mod Phys* **16**, 1

Some improvements in the use of relaxation methods for the solution of ordinary and partial differential equations

By L Fox, D.Phil., *The National Physical Laboratory*

(*Communicated by Sir Charles Darwin, F R S —Received* 11 *April* 1946)

In the solution by numerical methods of ordinary and partial differential equations, the derivatives are replaced by their finite-difference equivalents, usually infinite series of differences For the employment of relaxation methods it has been customary to ignore all but the dominant first term of these series, taking sufficient intervals to ensure the effective vanishing of the neglected terms In this paper, methods are developed whereby the full difference equations can be used at the maximum interval consistent with the convergence of the differences The number of mesh points, and hence the labour and difficulty of the relaxation process, is thus considerably reduced, and increased accuracy is obtained The method is applied to eight examples, including ordinary and partial differential equations, eigen-value problems, and the more difficult problem of curved boundaries

Introduction

The relaxation method, in an ultimate analysis, is concerned with the solution of simultaneous algebraic equations The equations may be known exactly, as in systems with finite degrees of freedom, or they may represent finite-difference approximations to a differential equation R V Southwell (1940) has shown in his book and in a series of papers published by the Royal Society (Southwell *et al* 1938-43), the power of relaxation methods in solving such equations Almost any given set of equations can be solved to any desired order of accuracy. For finite systems, the problem is then completely solved, but for differential equations it is also necessary to investigate the accuracy of the finite-difference approximations

This is a problem outside the scope of relaxation methods and for this reason has received only passing attention in Southwell's series, which has aimed chiefly at extending the use of relaxation methods to more and more problems of engineering interest For such problems high accuracy is not particularly necessary nor even desirable, but differential equations of similar type are of frequent occurrence in problems of pure and applied mathematics, and it is often very necessary either to have an accurate solution or to know the degree of accuracy of any approximate solution.

This paper makes full use of the theory of finite differences in an attempt to show how, in many cases, the accuracy attainable by relaxation methods can be brought comparable to that attained when, for example, standard step-by-step methods are used for the numerical integration of ordinary differential equations

Section 1 gives the basic finite-difference equations and indicates the methods of solution adopted. In § 2 examples are given of the solution of ordinary differential equations, the second of which is non-linear. Section 3 deals with two examples of

the solution of equations of the Poisson type, and § 4 brings eigen-value problems into the scope of the method. Finally § 5 gives examples of its application to more difficult problems involving curved boundaries. Of the eight examples treated, all but one have an analytical solution, so that the accuracy of the relaxation solutions is known exactly.

The differential equations considered in this paper are of the second order, and the only boundary condition is that in which the function is prescribed on a closed boundary. Treatment of other differential equations and boundary conditions, in particular the class in which some derivative is specified at a boundary, is reserved for a subsequent paper. The treatment of singularities is likewise omitted Any finite-difference method must necessarily break down in the neighbourhood of a singular point, and special methods have to be devised, depending on the type of the existing singularity

1. Finite-difference equations

The notation used conforms to what is now fairly standardized for finite differences The symbol Δ_r^n denotes the nth forward difference at the pivotal point r, δ_r^n the corresponding central difference at the same point, and the interval of tabulation is throughout denoted by the letter h. The suffix x or y means 'in the x-direction' or 'in the y-direction', and other symbols are self-explanatory.

Derivatives can be expressed in terms of differences by known formulae in the theory of finite differences (Tables 1942)

$$hf_0' = \delta_0' - \tfrac{1}{6}\delta_0''' + \tfrac{1}{30}\delta_0^{\mathrm{v}} - \ldots, \tag{1}$$

$$h^2 f_0'' = \delta_0'' - \tfrac{1}{12}\delta_0^{\mathrm{iv}} + \tfrac{1}{90}\delta_0^{\mathrm{vi}} - \ldots, \tag{2}$$

and similarly for higher orders The differences are arranged according to the following scheme.

$$\begin{array}{ccccc} f_{-1} & & \delta''_{-1} & & \delta^{\mathrm{iv}}_{-1} \\ & \delta'_{-\frac{1}{2}} & & \delta'''_{-\frac{1}{2}} & \\ f_0 & & \delta''_0 & & \delta^{\mathrm{iv}}_0 \quad \text{etc} \\ & \delta'_{\frac{1}{2}} & & \delta'''_{\frac{1}{2}} & \\ f_1 & & \delta''_1 & & \delta^{\mathrm{iv}}_1 \end{array} \qquad \begin{array}{l} \delta'_0 = \tfrac{1}{2}(\delta'_{-\frac{1}{2}} + \delta'_{\frac{1}{2}}) \\ \\ \qquad\qquad \text{etc} \\ \delta''_{\frac{1}{2}} = \tfrac{1}{2}(\delta''_0 + \delta''_1) \end{array}$$

For the application of relaxation methods it is convenient to write equations (1) and (2) in the alternative forms

$$hf_0' = \tfrac{1}{2}(f_1 - f_{-1}) - \tfrac{1}{6}\delta_0''' + \tfrac{1}{30}\delta_0^{\mathrm{v}} - \ldots, \tag{3}$$

$$h^2 f_0'' = (f_{-1} + f_1 - 2f_0) - \tfrac{1}{12}\delta_0^{\mathrm{iv}} + \tfrac{1}{90}\delta_0^{\mathrm{vi}} - \ldots, \tag{4}$$

etc

Laplace's operator in Cartesian co-ordinates x and y can then be expressed in the form

$$h^2\nabla^2 f_0 = (f_{-1} + f_1)_x + (f_{-1} + f_1)_y - 4f_0 - \tfrac{1}{12}(\delta^{\mathrm{iv}}_{0,x} + \delta^{\mathrm{iv}}_{0,y}) + \tfrac{1}{90}(\delta^{\mathrm{vi}}_{0,x} + \delta^{\mathrm{vi}}_{0,y}) - \ldots. \tag{5}$$

For the corresponding operator appropriate to problems of axial symmetry,

$$\left.\begin{aligned} h^2\left(\frac{\partial^2}{\partial r^2}+\frac{1}{r}\frac{\partial}{\partial r}+\frac{\partial^2}{\partial z^2}\right)f_0 &= (f_{-1}+f_1)_r+(f_{-1}+f_1)_z+\frac{h}{2r}(f_1-f_{-1})_r-4f_0+\Delta, \\ \text{where}\quad \Delta &= \frac{h}{r}\{-\tfrac{1}{6}\delta^{\prime\prime\prime}_{0,r}+\tfrac{1}{30}\delta^{\mathrm{v}}_{0,r}-\quad .\}-\tfrac{1}{12}(\delta^{\mathrm{iv}}_{0,r}+\delta^{\mathrm{iv}}_{0,z})+\tfrac{1}{90}(\delta^{\mathrm{vi}}_{0,r}+\delta^{\mathrm{vi}}_{0,z})-. \end{aligned}\right\} \quad (6)$$

Quantities such as Δ of equation (6) will in general be called the 'terms involving differences' or just 'the difference correction' The terms involving the functional values are called 'Lagrangian terms'

For problems in two variables it is not of course necessary to use the same interval in each variable The interval will usually depend on the shape of the boundary and the behaviour of the function, and may also be different in different parts of the field The use of a constant interval, however, makes for simplicity in the writing of the equations.

Southwell obtains (Christopherson & Southwell 1938) an equation analogous to (5) in the form

$$(f_{-1}+f_1)_x+(f_{-1}+f_1)_y-4f_0-h^2\nabla^2 f_0=\frac{h^4}{16}\nabla^4 f_0+\ldots,$$

in which coefficients of powers of h higher than the second are uncertain It can be shown that differences can be expressed in terms of derivatives (Tables 1942), the relevant equation being

$$\delta^{\prime\prime}_{0,x}=(f_{-1}+f_1)_x-2f_0=h^2f^{\prime\prime}_{0,x}+\tfrac{1}{12}h^4f^{\mathrm{iv}}_{0,x}+\tfrac{1}{360}h^6f^{\mathrm{vi}}_{0,x}+\quad ,$$

so that the correct form of Southwell's equation is given by

$$(f_{-1}+f_1)_x+(f_{-1}+f_1)_y-4f_0-h^2\nabla^2 f_0=\frac{h^4}{12}(f^{\mathrm{iv}}_{0,x}+f^{\mathrm{iv}}_{0,y})+\frac{h^6}{360}(f^{\mathrm{vi}}_{0,x}+f^{\mathrm{vi}}_{0,y})+\quad , \quad (7)$$

in which no terms of the form $\partial^4 f/\partial x^2\partial y^2$, etc , appear

In the use by relaxation methods of equations like (7) it has been customary to ignore the right-hand side of that equation, and to proceed by stages to such a small interval that the neglected terms were in fact negligible This has several disadvantages First, a very great number of points have often to be taken in the relaxation process to ensure the vanishing of the difference correction, and the difficulty and labour of obtaining an accurate solution to the simultaneous equations increases rapidly with the number of equations involved Secondly, derivatives of the computed function are often required in practical problems, and equations like (1) and (2) show that the nth derivative depends for its accuracy on the nth difference, which gets small with the interval of tabulation Thirdly, the criterion used initially for the legitimacy of neglect of difference corrections, namely that relaxation on successive meshes gives the same result, has tended to be ignored in later applications

The experienced user of step-by-step methods of integration will be familiar with at least the first two of the difficulties outlined. In general, the best interval to use is the one for which the differences of the function converge reasonably quickly The standpoint adopted in this paper follows this general rule, and equations like (5) are used to the full whenever they are convergent In practice it is found that fifth and higher differences can usually be ignored

The general method of solution is then as follows First, the difference corrections are ignored, and a first approximation obtained by relaxation methods in the usual way. These values are then differenced, and, if the differences are obviously converging, the difference corrections are calculated and entered as a new residual, to be relaxed in turn. This is continued until the full finite-difference equations are satisfied The actual details depend on the type of differential equation considered, but for linear equations the procedure can be expressed mathematically as follows Let

$$L(f)+\Delta(f)+F=0$$

be the exact finite-difference equivalent of the differential equation, where $L(f)$ denotes the Lagrangian terms, $\Delta(f)$ the difference correction, and F any constant or function of the independent variables Let $f^{(0)}$ take the boundary values of the required solution f, and satisfy the equation

$$L(f^{(0)})+F=0.$$

Successive corrections $f^{(1)}, f^{(2)}, \ldots$ are then obtained by solving the equations

$$L(f^{(1)})+\Delta(f^{(0)})=0,$$
$$L(f^{(2)})+\Delta(f^{(1)})=0,$$

etc , in which $f^{(1)}, f^{(2)}, \ldots$ have boundary values zero. In practice the successive corrections usually tend rapidly to zero

In eigen-value problems or in problems involving non-linear differential equations or boundary conditions solutions cannot be superposed in this way, but the required modifications are fairly obvious. The essential requirement is that the final solution, however it is obtained, should satisfy the full finite-difference equations at every point.

This method of solution, when it can be used, goes far towards removing the three disadvantages quoted, and gives greater accuracy with less labour. Some examples of its application are given in the following sections

2. Ordinary differential equations

The solution of an ordinary differential equation of the second order and first degree is in general completely defined when two boundary conditions are specified. The common types of boundary condition are those in which either the function and its first derivative are known at one end of the range of integration, or the value of the function is specified at each end of the range. Relaxation methods are needed

only for the second type of boundary condition. This is appropriate since step-by-step methods cannot be applied directly to such problems, but work very well for problems governed by the first type of boundary condition.

Example 1

A simple illustration of the methods of the last section is provided by the differential equation

$$2d^2f/dx^2 + df/dx + 100\sin x = 0 \tag{8}$$

with the boundary conditions

$$f = 0 \quad \text{at} \quad x = 0,\ 10.$$

Application of equations (3) and (4) leads to the finite-difference equation

$$\left.\begin{aligned} \left(2+\frac{h}{2}\right)f_1 + \left(2-\frac{h}{2}\right)f_{-1} - 4f_0 + 100h^2\sin x + \Delta = 0, \\ \text{where} \qquad \Delta = -\frac{h}{6}\delta_0''' - \tfrac{1}{8}\delta_0^{\text{IV}} + \frac{h}{30}\delta_0^{\text{V}} + \tfrac{1}{56}\delta_0^{\text{VI}} - \quad . \end{aligned}\right\} \tag{9}$$

Taking an interval $h = 1$ and ignoring the difference correction, a first approximation $f^{(0)}$ is obtained by relaxation.

Table 1 gives the values of $f^{(0)}$ and its first four differences The difference correction $\Delta(f^{(0)})$, also given in table 1, can be computed directly at every point except $x = 1$ and $x = 9$, at which points the fourth differences are not known The required differences are obtained reasonably accurately by extrapolation (by plotting or otherwise), and the corresponding values of $\Delta(f^{(0)})$ can then be obtained The way in which these extrapolated values enter into the difference correction shows that fairly large inaccuracies are tolerable in the estimates. All extrapolated values are given in brackets in table 1.

TABLE 1

x	$f^{(0)}$	δ'	δ''	δ'''	δ^{IV}	$\Delta(f^{(0)})$
0	0		—		—	—
		+53		(−10)		
1	+53		−55		(+30)	(− 6)
		− 2		+20		
2	+51		−35		+24	− 0
		−37		+44		
3	+14		+ 9		−11	− 5
		−28		+33		
4	−14		+42		−42	+ 5
		+14		− 9		
5	0		+33		−32	+10
		+47		−41		
6	+47		− 8		+ 7	+ 5
		+39		−34		
7	+86		−42		+37	− 4
		− 3		+ 3		
8	+83		−39		+37	−10
		−42		+40		
9	+41		+ 1		(0)	(−7)
		−41		(+40)		
10	0		—		—	—

As a check on the extrapolation, or as an alternative method of computation, Δ may itself be plotted to obtain the required values at the points adjacent to the two ends of the range The combined effect of fifth and higher differences in the computation of Δ is small enough to be ignored

The first correction $f^{(1)}$, which satisfies the equations

$$2{\cdot}5f_1^{(1)} + 1{\cdot}5f_{-1}^{(1)} - 4f_0^{(1)} + \Lambda(f_0^{(0)}) = 0,$$

is then obtained, and its values are given in table 2.

TABLE 2

x	0	1	2	3	4	5	6	7	8	9	10
$f^{(1)}$	0	−5	−6	−3	+1	+1	−3	−7	−8	−5	0

Further corrections are found to be negligible, and the solution is therefore given by the sum of $f^{(0)}$ and $f^{(1)}$ It is desirable, however, to test this solution in equation (9), since the accumulation of unliquidated residuals in the partial solutions may affect the correct rounding-off at some points

TABLE 3

x	f	δ'	δ''	δ'''	δ^{iv}	$\Lambda(f)$	R	analytical solution
0	0		—	(−10)	—	—	—	0
		+48		+20	(+30)	(−6)		
1	+48		−51				−1	+47 7
		− 3		+41	+21	−9		
2	+45		−31				+2	+45·3
		−34		+28	−13	−4		
3	+11		+10				+1	+11 6
		−24		− 9	−37	+5		
4	−13		+38				0	−12 4
		+14		−37	−28	+9		
5	+ 1		+29				−1	+ 1 4
		+43		−31	+ 6	+5		
6	+44		− 8				0	+44 0
		+35		+ 4	+35	−4		
7	+70		−39				0	+78 5
		− 4		+38	+34	−9		
8	+75		−35				−2	+74 5
		−39		(+38)	(0)	(−6)		
9	+36		+ 3				+3	+36 5
		−36			—	—		
10	0		—				—	0

The sum f of the partial solutions $f^{(0)}$ and $f^{(1)}$ is shown in table 3, together with its differences, the computed difference correction, and the final residuals given by equation (9) These residuals indicate that the values of the function should be slightly increased (by less than one unit) at the points $x = 2$, 3 and 9, and slightly decreased at $x = 8$ The computed values of the analytical solution

$$f = 38\ 94 - 58{\cdot}94e^{-\frac{1}{2}x} + 40\sin x + 20\cos x,$$

given on the right of table 3, bear out this conjecture, and show that the maximum error in the final result is less than 1 unit, compared with a maximum of 8 units when the difference correction is ignored. Since only two figures were taken in the computation, the increased accuracy is important

Example 2

The second example is taken from a problem submitted to the author by Mr Gainsborough, of the Radio Division, National Physical Laboratory. In experiments connected with the electrical heating of a wire it was desired to obtain solutions of the differential equation

$$d^2T/dx^2 = 1{\cdot}815 \times 10^{-12}T^5 - 3\ 272 \times 10^4 HT + 3{\cdot}742 \times 10^6 H - 4{\cdot}411. \qquad (10)$$

Plots were required of the temperature T for various values of the current H, the temperature being maintained at a constant value of 300° C at the two ends $x = 0$ and 1 of the wire Solutions can only be obtained with difficulty by step-by-step methods, and the time taken is of the order of five times that necessary for solution by relaxation methods.

The finite-difference equation in this case becomes

$$\left.\begin{aligned} &T_1 + T_{-1} - 2T_0 - h^2\{1{\cdot}815 \times 10^{-12} T_0^5 - 3{\cdot}272 \times 10^4 H T_0 + 3\,742 \times 10^6 H - 4\,411\} + \Delta = 0, \\ &\text{where} \qquad \Delta = -\tfrac{1}{12}\delta_0^{\text{iv}} + \tfrac{1}{90}\delta_0^{\text{vi}} - \end{aligned}\right\} \tag{11}$$

Here the relaxational technique is rather unusual, in that the change in T required to liquidate a residual at any point depends on the value of T at that point. A small change δT_0 affects the residuals at adjacent points by $+\delta T_0$, and at the point in question by an amount

$$\delta R_0 = \delta T_0[-2 - h^2\{1\,815 \times 10^{-12} \times 5T_0^4 - 3\,272 \times 10^4 H\}].$$

This equation does not hold precisely at the start of the relaxation process, when changes in T may be large, but the experienced relaxer will have little difficulty in securing rapid convergence

By symmetry only one-half of the range needs consideration, and the first two solutions obtained are shown in table 4

TABLE 4

	$H = 2 \times 10^{-4}$					$H = 3 \times 10^{-4}$				
x	T	δ'	δ''	δ'''	δ^{iv}	T	δ'	δ''	δ'''	δ^{iv}
0·0	300		—	—	—	300		—	—	—
0·1	434	+134	−21			576	+276	−44		
0·2	547	+113	−28	−7	+3	808	+232	−61	−17	+10
0·3	632	+85	−32	−4	+1	979	+171	−68	−7	+6
0·4	685	+53	−35	−3	+2	1082	+103	−69	−1	+2
0·5	703	+18	−36	−1	+2	1116	+34	−68	+1	−2

The differences of the functions indicate that the difference correction is negligible in the first case, and would affect the second solution only very slightly.

As H increases, the fourth differences become more significant For $H = 10^{-3}$, the first approximation $T^{(0)}$ and its differences are shown in table 5.

TABLE 5

x	$T^{(0)}$	δ'	δ''	δ'''	δ^{iv}	$\Delta(T^{(0)})$
0·0	300		(−61)		—	—
0·1	1119	+819	−296	(−235)	(+248)	(−21)
0·2	1642	+523	−283	+13	+119	−10
0·3	1882	+240	−151	+132	−48	+4
0·4	1971	+80	−67	+84	−61	+5
0·5	1993	+22	−44	+23	−46	+4

For all points except $x = 0{\cdot}1$, the difference correction $\Delta(T^{(0)})$, given on the right of table 5, is immediately calculable. At $x = 0{\cdot}1$ the fourth difference δ^{iv}_{01} is difficult to extrapolate, and an alternative method is used. The required fourth difference is built up from the second difference at $x = 0{\cdot}0$, which is first obtained from the second derivative at $x = 0{\cdot}0$, immediately calculable from the differential equation, in the form

$$\delta''_0 = h^2 T''_0 + \tfrac{1}{12}\delta^{iv}_0 - \ .$$

in which all but the first term on the right is neglected The second difference thus obtained, and the third and fourth differences built up from it, are shown in brackets in table 5, as is also the computed difference correction at $x = 0{\cdot}1$ It is clear that this process entails no very great error

When the difference corrections $\Delta(T^{(0)})$ are entered as a new residual, the solution is changed to that given in table 6 Further corrections are found to be negligible.

TABLE 6

x	0 0	0 1	0 2	0 3	0·4	0 5
T	300	1102	1629	1878	1972	1995

This solution is not completely reliable, since sixth and higher differences, not completely negligible, were ignored in the computation of the difference correction. From work done at smaller intervals, however, it is found that the maximum error in this solution is only 2 units, compared with an error of 20 units when the difference correction is ignored at this interval.

When H is increased still further to 5×10^{-3}, differences at an interval of 0·1 become divergent, and the interval 0 05 has to be taken Table 7 gives the first approximation $T^{(0)}$ and its differences, the computed values of $\Delta(T^{(0)})$, and the corrected solution T.

TABLE 7

x	$T^{(0)}$	δ'	δ''	δ'''	δ^{iv}	$\Delta(T^{(0)})$	T
0·00	300		(− 76)		—	—	300
		+1424		(−514)			
0 05	1724		−590		(+602)	−50	1687
		+ 834		+ 88			
0 10	2558		−502	+281	+193	−16	2535
		+ 332					
0·15	2890		−221	+145	−136	+11	2886
		+ 111					
0 20	3001		− 76	+ 52	− 93	+ 8	3002
		+ 35					
0 25	3036		− 24	+ 16	− 36	+ 3	3038
		+ 11					
0·30	3047		− 8	+ 6	− 10	+ 1	3048
		+ 3					
0 35	3050		− 2	+ 2	− 4	0	3051
		+ 1					
0 40	3051		0	− 1	− 3	0	3052
		+ 1					
0·45	3052		− 1	+ 1	+ 2	0	3052
		0					
0·50	3052		0		− 2	0	3052

The difference correction is not negligible until the interval is reduced by a factor 4, and the result obtained from work done at this interval nowhere differs by more than 1 unit from the solution given on the right of table 7. The application

of the difference correction on the coarse mesh thus reduces the maximum error from 37 units to 1 unit, and the saving of labour and the accuracy obtainable for non-linear equations of this type is clearly demonstrated.

3. Partial differential equations

The use of the difference correction takes on additional importance when relaxation methods are used for the solution of partial differential equations. It is here most important to use as large an interval as possible, for the number of equations is multiplied by four when the interval is halved, and the difficulty of obtaining accurate solutions is correspondingly increased.

Example 3

Consider, as a simple example, the differential equation

$$\nabla^2 w = \text{const.}, \tag{12}$$

with the boundary condition $w = 0$ on the lines $x = \pm 1$, $y = \pm 1$. By symmetry only one semi-quadrant of the square need be considered, and it will be convenient to take the constant in equation (12) to be -2560.

The finite-difference equation is

$$\left.\begin{aligned} &(w_1 + w_{-1})_x + (w_1 + w_{-1})_y - 4w_0 + 2560h^2 + \Delta = 0, \\ \text{where} \quad &\Delta = -\tfrac{1}{12}(\delta^{4}_{0,x} + \delta^{4}_{0,y}) + \tfrac{1}{90}(\delta^{6}_{0,x} + \delta^{6}_{0,y}) - \ldots. \end{aligned}\right\} \tag{13}$$

The method of solution is that outlined in § 1 Solutions are obtained successively to the sets of equations

$$\left.\begin{aligned} &(w_1^{(0)} + w_{-1}^{(0)})_x + (w_1^{(0)} + w_{-1}^{(0)})_y - 4w_0^{(0)} + 2560h^2 = 0, \\ &(w_1^{(1)} + w_{-1}^{(1)})_x + (w_1^{(1)} + w_{-1}^{(1)})_y - 4w_0^{(1)} + \Delta(w_0^{(0)}) = 0, \\ &\qquad\qquad \text{etc.} \end{aligned}\right\} \tag{14}$$

where, at the boundary,

$$w^{(0)} = w(=0), \quad w^{(1)} = w^{(2)} = \quad = 0$$

and the final solution is given by

$$w = w^{(0)} + w^{(1)} + \ldots \tag{15}$$

With an interval $h = 1$, there is only one point inside the field, and the result is as shown in figure 1 Here no correction for differences can be made, since Δ is incalculable.

For the smaller interval $h = \frac{1}{2}$, the first approximation $w^{(0)}$, obtained from the first of equations (14) is shown in figure 2.

When these values are differenced (there are only two independent lines of points) only one fourth difference can be obtained on each line, and as usual it is necessary to extrapolate this difference for points adjacent to the boundary One assumption is that the fourth difference is constant, but this can be improved upon by a method similar to that used in example 2 This consists of applying the differen-

tial equation at points on the boundary, at which the second derivative in one direction is known (zero in this case) The other second derivative can thus be calculated, and from it a reasonably accurate assumption for the second difference at the boundary From this the fourth difference at points adjacent to the boundary can be built up. The differences, with extrapolated values given in brackets, are shown in table 8.

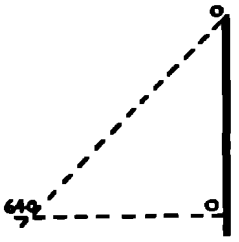


FIGURE 1

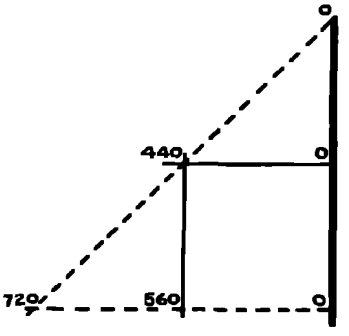


FIGURE 2 $w^{(0)}$.

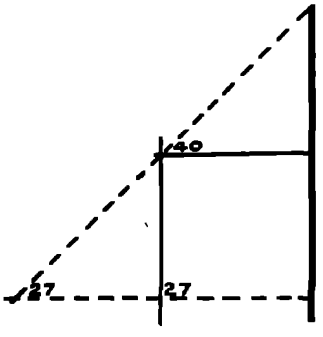


FIGURE 3. $\Delta(w^{(0)})$

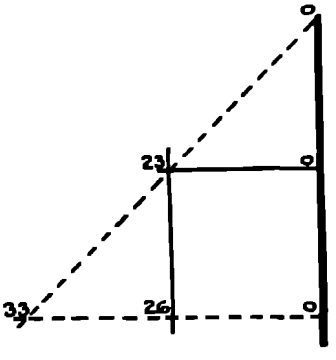


FIGURE 4. $w^{(1)}$.

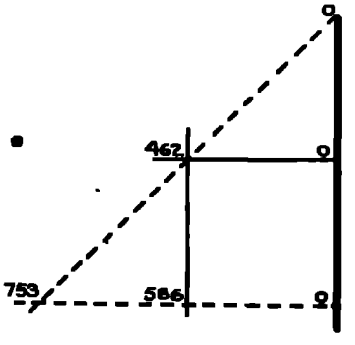


FIGURE 5. w

TABLE 8

$w^{(0)}$	δ'	δ''	δ'''	δ^{iv}	$w^{(0)}$	δ'	δ''	δ'''	δ^{iv}
0		(−640)		—	0		(−640)		—
	+440		(+320)			+560		(+240)	
440		−320		(−240)	560		−400		(−160)
	+120		+ 80			+160		+ 80	
560		−240		−160	720		−320		−160
	−120		− 80			−160		− 80	
440		−320		(−240)	560		−400		(−160)
	−440		(−320)			−560		(−240)	
0		(−640)		—	0		(−640)		—

The difference correction $\Delta(w^{(0)})$ can then be computed, and used for the calculation by relaxation of a correction $w^{(1)}$. Values of $\Delta(w^{(0)})$ and of $w^{(1)}$ are shown in figures 3 and 4 respectively Further corrections are almost negligible, and the final solution, adjusted slightly to allow for the accumulation of residuals, is given in figure 5.

The solution of figure 5 is not completely reliable, since sixth differences were ignored in the computation of the difference correction, which has in addition small

errors at points adjacent to the boundary due to the extrapolation described above Comparison with the analytical solution given below (figure 10) shows, however, that its maximum error is only one unit, compared with a maximum of 34 when the difference correction is neglected

The difference correction is not negligible even at the smaller interval $h = \frac{1}{4}$ Figure 6 gives the first approximation $w^{(0)}$ at this interval.

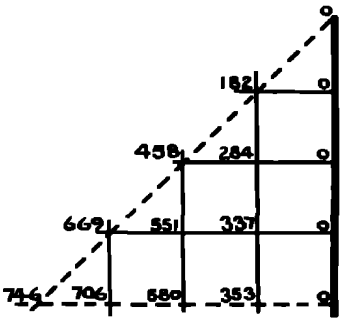


FIGURE 6. $w^{(0)}$.

When these values are differenced, as shown in table 9, enough fourth differences are obtainable to enable the required extrapolation to be made without the special device previously used. Extrapolated values are as usual given in brackets.

TABLE 9

$w^{(0)}$	δ'	δ''	δ'''	δ^{iv}	$w^{(0)}$	δ'	δ''	δ'''	δ^{iv}
0		—	—	—	0		—	—	—
	+353					+337			
353		−126		(−12)	337		−123		(−12)
	+227		+25			+214		+27	
580		−101		−10	551		− 96		−12
	+126		+15			+118		+15	
706		− 86		− 9	669		− 81		− 8
	+ 40		+ 6			+ 37		+ 7	
746		− 80		−12	706		− 74		−14
0		—	—	—	0		—	—	—
	+284					+182			
284		−110		(−12)	182		− 80		(−24)
	+174		+29			+102		+31	
458		− 81		−12	284		− 49		−19
	+ 93		+17			+ 53		+12	
551		− 64		−11	337		− 37		− 7
	+ 29		+ 6			+ 16		+ 5	
580		− 58		−12	353		− 32		−10

The difference correction $\Delta(w^{(0)})$ is small, but its combined effect is large enough to cause a significant correction $w^{(1)}$ These values are given respectively in figures 7 and 8

Further corrections are negligible, apart from small adjustments due to accumulation of small residuals, and the final solution is shown in figure 9 For comparison the analytical solution (Timoshenko 1934, p 245), given in this case by

$$w = \frac{32}{\pi^3} \cdot 1280 \sum_0^\infty \frac{(-1)^n}{(2n+1)^3}\left(1 - \frac{\cosh(2n+1)\pi y/2}{\cosh(2n+1)\pi/2}\right)\cos(2n+1)\pi x/2$$

with the origin at the mid-point of the square, was computed and given in figure 10.

A still finer mesh is needed before the difference correction can be ignored, and the above results indicate clearly the advantage of the use of the difference correction on as coarse a mesh as possible, both as regards accuracy and labour.

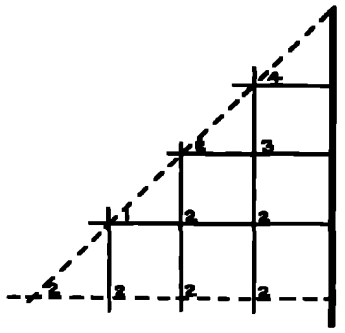


FIGURE 7. $\Delta(w^{(0)})$

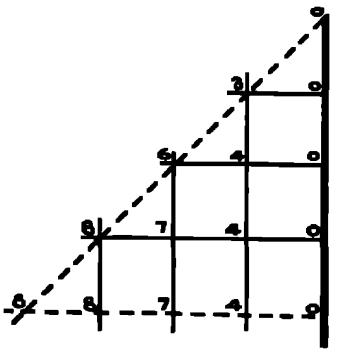


FIGURE 8 $w^{(1)}$

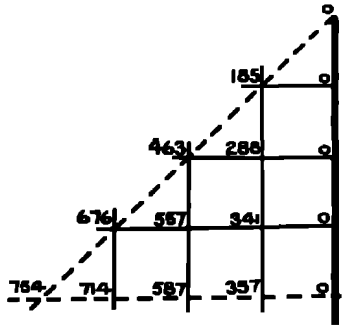


FIGURE 9 Final solution.

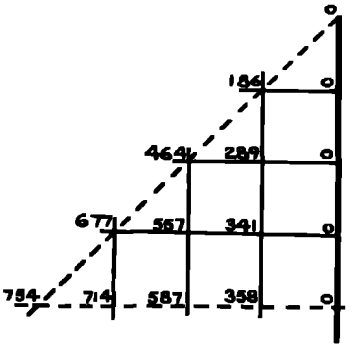


FIGURE 10. Analytical solution

Example 4

When the differential equation contains a first derivative, the inclusion of the difference correction is even more important, for it includes differences of the third as well as of higher orders This is illustrated by the next example, which was treated in Allen, Fox & Southwell (1943)

The equation to be solved is

$$\frac{\partial^2\phi}{\partial r^2}-\frac{3}{r}\frac{\partial\phi}{\partial r}+\frac{\partial^2\phi}{\partial z^2}+25600=0, \tag{16}$$

with the boundary conditions $\phi = 0$ on the lines $r = 1, 2, z = \pm\frac{1}{2}$ (figure 11).

Combination of equations (3) and (4) leads to the finite-difference equation

$$\left.\begin{aligned}&\left\{\left(1-\frac{3h}{2r_0}\right)\phi_1+\left(1+\frac{3h}{2r_0}\right)\phi_{-1}\right\}_r+(\phi_1+\phi_{-1})_z-4\phi_0+25600h^2+\Delta=0,\\ &\text{where}\qquad \Delta=\frac{h}{2r_0}\delta^{iv}_{0,r}-\tfrac{1}{12}(\delta^{iv}_{0,r}+\delta^{iv}_{0,z})-\ldots.\end{aligned}\right\} \tag{17}$$

With the difference correction neglected, the first approximation $\phi^{(0)}$ for a mesh of length $h = \frac{1}{4}$ is shown in figure 12.

When these values are differenced with a view to the computation of the difference correction, the usual extrapolation is necessary for points adjacent to the boundary The method used is that of the last example Satisfaction at boundary points of the first of equations (17) in which A is neglected, gives a reasonably accurate estimate for points on mesh lines external to the boundary, and from these the required differences can be obtained In table 10 the extrapolated values are in brackets

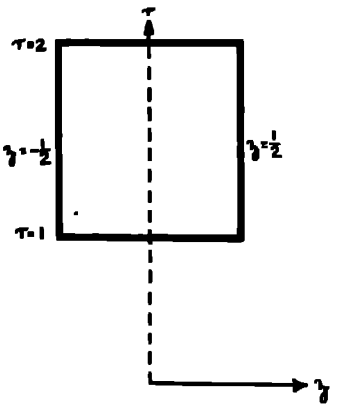


FIGURE 11

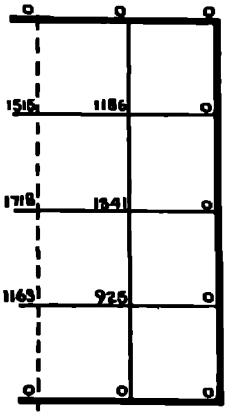


FIGURE 12. $\phi^{(0)}$

TABLE 10

$z = 0$

r	$\phi^{(0)}$	δ'	δ''	δ'''	δ^{iv}
(0 75)	(−1692)		—	—	—
1 00	0	(+1692)	(−529)		—
1·25	+1163	+1163	− 608	(−79)	(− 71)
1 50	+1718	+ 555	− 758	− 150	−404
1·75	+1515	− 203	−1312	− 554	(−802)
2 00	0	−1515	(−2668)	(−1356)	—
(2 25)	(−4183)	(−4183)	—	—	—

$z = 0\ 25$

r	$\phi^{(0)}$	δ'	δ''	δ'''	δ^{iv}
	(−1584)		—	—	—
	0	(+1584)	(− 659)		—
	+ 925	+ 925	− 509	(+ 150)	(− 212)
	+1341	+ 416	− 571	− 62	− 398
	+1186	− 155	−1031	− 460	(−1026)
	0	−1186	(−2517)	(−1486)	—
	(−3703)	(−3703)	—	—	—

$r = 1{\cdot}25$

z	$\phi^{(0)}$	δ'	δ''	δ'''	δ^{iv}
0·00	+1163		− 476		−422
0 25	+ 925	− 238	− 687	−211	(−702)
0 50	0	− 925	(−1600)	(−913)	—
(0·75)	(−2525)	(−2525)	—	—	—

$r = 1\ 50$

z	$\phi^{(0)}$	δ'	δ''	δ'''	δ^{iv}
	+1718		− 754		−420
	+1341	− 377	− 964	−210	(−426)
	0	−1341	(−1600)	(−636)	—
	(−2941)	(−2941)	—	—	—

$r = 1{\cdot}75$

z	$\phi^{(0)}$	δ'	δ''	δ'''	δ^{iv}
0 00	+1515		− 658		−398
0·25	+1186	− 329	− 857	−199	(−544)
0·50	0	−1186	(−1600)	(−743)	—
(0 75)	−2786	(−2786)	—	—	—

Calculation of the difference correction $\Delta(\phi^{(0)})$ shown in figure 13 is then straightforward, and the correction $\phi^{(1)}$ (figure 14) follows by relaxation.

In this case the difference correction $\Delta(\phi^{(1)})$ is not negligible, and a further correction solution $\phi^{(2)}$ can be obtained These quantities are given in figures 15 and 16. The final solution, given by the sum of the three partial solutions and adjusted slightly for the accumulation of residuals, is shown in figure 17

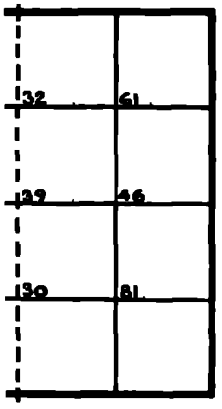


FIGURE 13. $\Delta(\phi^{(0)})$

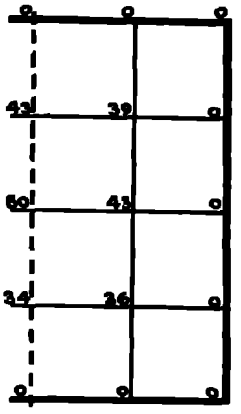


FIGURE 14 $\phi^{(1)}$

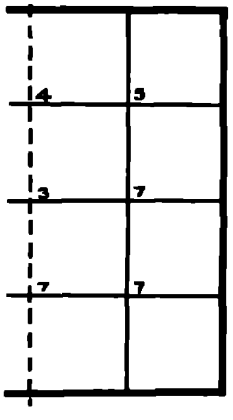


FIGURE 15. $\Delta(\phi^{(1)})$.

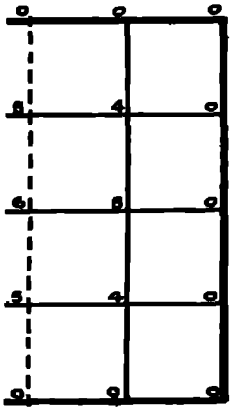


FIGURE 16 $\phi^{(2)}$.

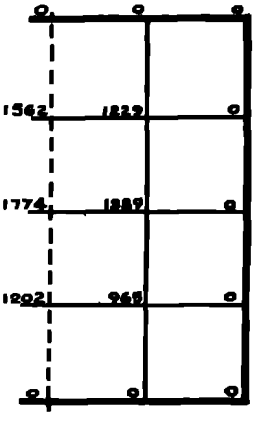


FIGURE 17. ϕ

Although the result of figure 17 is not completely reliable, due to the extrapolations and the neglect of fifth and higher differences in the computation of the difference corrections, comparison with the analytical solution given below (figure 22) shows that its maximum error is only 2 units, compared with a maximum of 56 when the difference correction is ignored

On the next mesh $h = \frac{1}{8}$, the values of $\phi^{(0)}$, $\Delta(\phi^{(0)})$ and $\phi^{(1)}$ are shown in figures 18, 19 and 20. Here the special device previously used for the computation of Δ at points adjacent to the boundary is unnecessary

Further corrections are negligible, except for the usual slight adjustments, and the final solution is given in figure 21. For comparison the analytical solution has been computed and the results are given with an extra figure in figure 22 This solution, due to Michell (1900), obtains the stress-function ϕ in the form

$$\phi = 12800(z^2 - \tfrac{1}{4}) + \phi',$$

$$\phi' = r^2 \Sigma \{C_n I_2(n\pi r) + D_n K_2(n\pi r)\} \sin n\pi(z + \tfrac{1}{2}),$$

where n is an integer and I_2, K_2 Bessel functions in the usual notation.

The coefficients are determined from the equations

$$\left.\begin{aligned} C_{2n} = D_{2n} &= 0, \\ C_{2n+1} I_2\{(2n+1)\pi\} + D_{2n+1} K_2\{(2n+1)\pi\} &= -102400/(2n+1)^3\pi^3, \\ C_{2n+1} I_2\{(2n+1)2\pi\} + D_{2n+1} K_2\{(2n+1)2\pi\} &= -25600/(2n+1)^3\pi^3 \end{aligned}\right\}$$

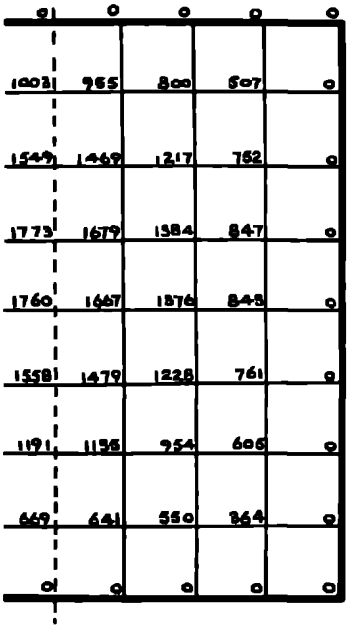


FIGURE 18 ϕ^0

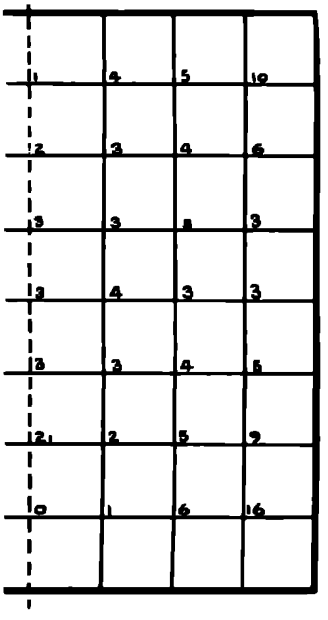


FIGURE 19 $\Delta(\phi^{(0)})$

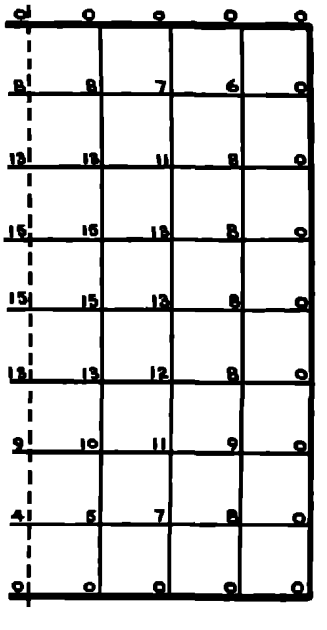


FIGURE 20 $\phi^{(1)}$

Comparison shows that the maximum error is 1 unit except at the single point $r = \frac{5}{8}$, $z = \frac{3}{8}$, where the error is 2 units, whilst the errors involved in neglect of the difference correction have a maximum of 15 units and an average of 10 units.

It is also instructive to compare these results with that of D. N Allen, who performed the corresponding computations (Allen *et al* 1943) The relaxation solution is there given to one extra figure on a finer mesh $h = \frac{1}{16}$, but it is found that only the first three figures are accurate, the maximum error being about 40 units in the fifth figure The difference correction in fact is not entirely negligible even to four figures on this fine mesh

Allen's remark that the extra figures, though inaccurate, are useful for the determination of derivatives, is true for points remote from the boundary, but obviously breaks down near the boundary, where the stresses are greatest and most important. It is clear that the use of the difference correction on as coarse a mesh as possible improves the accuracy and reduces considerably the labour.

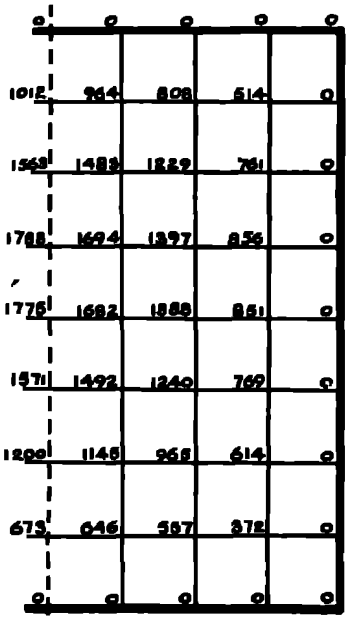


FIGURE 21. Final solution

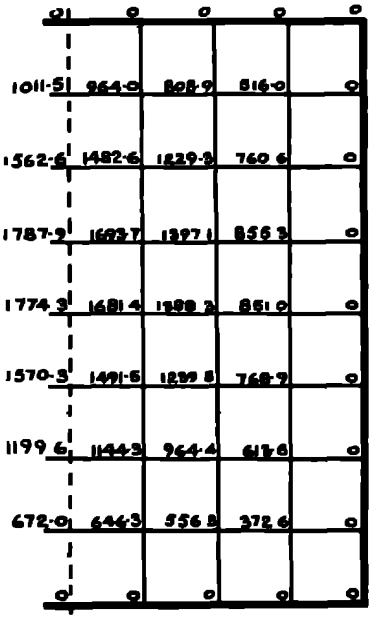


FIGURE 22. Analytical solution.

4 EIGEN-VALUE PROBLEMS

The application of the difference correction in eigen-value problems presents no extra difficulty. In Allen *et al* (1942) details are given of the relaxational treatment of such problems, and it remains only to show how an approximate solution, obtained with neglect of the difference correction, can be improved by its inclusion.

The equation governing the free vibrations of a membrane is given by

$$\nabla^2 w + \lambda w = 0, \tag{18}$$

the function vanishing at the boundary.

The relaxation method starts with an approximation (not necessarily very close) to a mode w, from which is calculated, by Rayleigh's principle, a first estimate for the corresponding eigen-value in the form

$$\lambda = -\frac{\iint w . \nabla^2 w \, dx\, dy}{\iint w^2 \, dx\, dy}. \tag{19}$$

Relaxation is then used to reduce as far as possible the residuals calculated from the finite-difference equations, in which the computed value of λ is used. When no further improvement is possible, a new value of λ is calculated by equation (19), and the mode and eigen-value are continuously improved by iteration

By this means a first approximation $w^{(0)}$ to the mode and $\lambda^{(0)}$ to the eigen-value are obtained They satisfy equations (18) and (19) in which $\nabla^2 w$ is replaced by the approximate finite-difference expression

$$h^2\nabla^2 w_0 = (w_1 + w_{-1})_x + (w_1 + w_{-1})_y - 4w_0$$

This solution is then improved by the following sequence of operations

(i) Calculate the difference corrections

$$\Delta(w_0^{(0)}) = -\tfrac{1}{12}(\delta^{iv}_{0,x} + \delta^{iv}_{0,y}) + \tfrac{1}{90}(\delta^{vi}_{0,x} + \delta^{vi}_{0,y}) - \quad ..$$

(ii) Calculate, at each mesh point, a more accurate value for $\nabla^2 w^{(0)}$, given by equation (5) of § 1.

(iii) Using these values, compute a new value of λ from equation (19).

(iv) Insert this value in the accurate finite-difference equation

$$(w_1 + w_{\ 1})_x + (w_1 + w_{-1})_y - 4w_0 + \lambda h^2 w_0 + \Delta = 0, \tag{20}$$

calculate new residuals, and relax if necessary.

(v) Repeat until λ and w satisfy the accurate finite-difference equation (20) and (necessarily), Rayleigh's principle (equation (19))

It is fairly obvious from the form of the differential equation that the difference correction will in general be roughly proportional to the functional value, and hence the correction to the eigen-value will be far more significant than the correction to the mode This is found to be the case in the examples which follow

Example 5

A first example is provided by the determination of the mode and eigen-value of the fundamental vibration of a square membrane, the first problem treated in Allen *et al.* (1942)

If the side of the membrane has unit length, the analytical solution is known to be

$$\left.\begin{aligned} w &= \cos\pi x \cos\pi y, \\ \lambda &= 2\pi^2 = 19{\cdot}74. \end{aligned}\right\} \tag{21}$$

With a mesh of length $h = \frac{1}{4}$, an approximation $w^{(0)}$ is first obtained in the usual way. This solution is given in figure 23, in which only one semi-quadrant of the square need be given. The figures on the right of the nodal points are the values of $\nabla^2 w^{(0)}$, computed with neglect of the difference correction The corresponding approximation to the eigen-value is found to be

$$h^2\lambda^{(0)} = 1\ 172 \quad \text{or} \quad \lambda^{(0)} = 18{\cdot}75.$$

When these values are differenced, estimates for the fourth differences at points adjacent to the boundary are obtained by the method used in example 3 Table 11 gives the differences, with extrapolated values in brackets

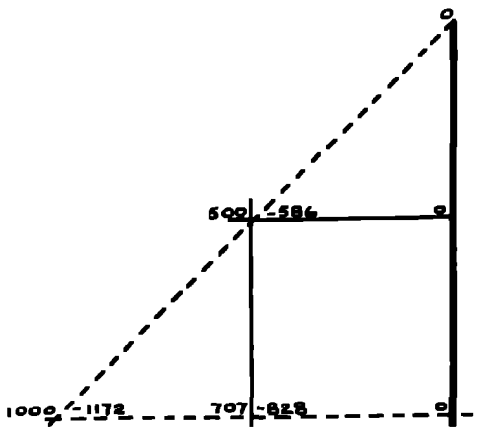


FIGURE 23 $w^{(0)}$ and $\nabla^2 w^{(0)}$

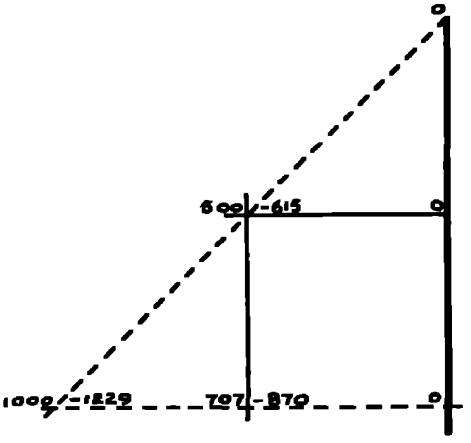


FIGURE 24. w and $\nabla^2 w$

TABLE 11

$w^{(0)}$	δ'	δ''	δ'''	δ^{iv}	$w^{(0)}$	δ'	δ''	δ'''	δ^{iv}
(−500)		—	—	—	(−707)		—	--	--
0	(+500)				0	(+707)			
+500	+500	(0)		—	+707	+707	(0)	(−414)	-
+707	+207	−293	(−293)	+172	+1000	+293	−414	−172	(+242)
		−414	−121	+242			−586		+344

When the difference correction is computed, using only the fourth difference, the values of $\nabla^2 w$ are all greater than those given by the approximate formula, and the new eigen-value is found to be

$$h^2\lambda = 1\ 229 \quad \text{or} \quad \lambda = 19\ 66.$$

The resulting residuals are found to be negligible, and hence the approximate solution $w^{(0)}$ is unchanged It does, in fact, agree accurately with the analytical solution. Figure 24 gives the final solution, together with the more accurate values of $\nabla^2 w$.

The computed eigen-value has an error of 0·4 %, compared with 5 % when the difference correction is ignored The only source of error in the final solution is the neglect of sixth differences and the extrapolation at the boundary

In the treatment of the same problem in Allen *et al* (1942) (for which the computations were performed by the present writer) a value $\lambda = 19\ 652$ was obtained on a mesh twice as fine as that of figure 24 It was remarked at the time that the application of Rayleigh's principle gave a value for λ which was less than the correct result, contradicting the theorem that, for the gravest mode, an overestimate will be obtained This was put down to the use of approximate quadrature formulae in the

double integration, but it is now clear that the main source of error was the neglect of fourth and higher differences in the computation of the Laplacian

Example 6

For higher modes, in which there are one or more nodal lines, a finer mesh must be used to secure reasonable convergence of the differences. In this example a higher symmetrical mode is obtained, the analytical solution being

$$\left.\begin{aligned} w &= \cos 3\pi x \cos \pi y + \cos 3\pi y \cos \pi x, \\ \lambda &= 10\pi^2 = 98\ 70 \end{aligned}\right\} \tag{22}$$

With a mesh of interval 0·1, the first approximate solution is shown in figure 25, the computed eigen-value being

$$h^2\lambda^{(0)} = 0\ 922 \quad \text{or} \quad \lambda^{(0)} = 92\ 2$$

In figure 25 the approximate values of $\nabla^2(w^{(0)})$ are given on the right of the nodal points.

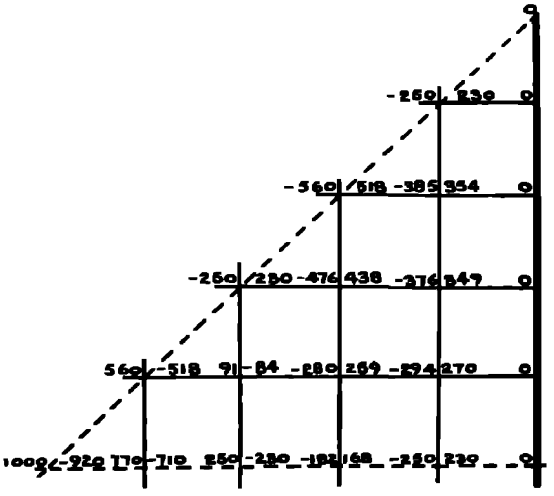


FIGURE 25 $w^{(0)}$ and $\nabla^2 w^{(0)}$ approximately.

When these values are differenced, one extrapolation is as usual necessary for the determination of the fourth difference at points adjacent to the boundary Using fourth differences only, the difference correction to $\nabla^2(w^{(0)})$ is given in figure 26 Applying Rayleigh's principle to these values, a correction is obtained,

$$\delta\lambda^{(0)} = 5\ 8,$$

to the previously computed value $\lambda^{(0)}$ The new residuals are found to be zero, so that the modal values are unchanged.

The ignored sixth differences are, however, significant, and they make a contribution to $\nabla^2 w$ which is shown in figure 27. The determination of this correction

for points adjacent to the boundary would involve yet another extrapolation, and so no values are given at these points.

These values are again roughly proportional to the modal values, and a further application of Rayleigh's principle results in a second correction to the eigen-value,

$$\delta\lambda = 0\ 6.$$

The final value is thus given by $\lambda = 98{\cdot}6$,

agreeing very closely with the analytical result. The mode has not changed from its first approximation $u^{(0)}$, which, in fact, differs from the analytical solution by less than 1 unit

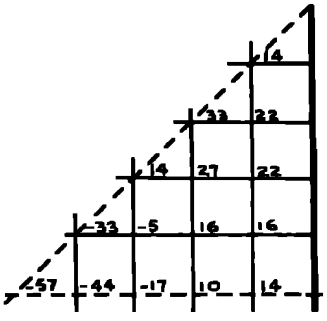


FIGURE 26

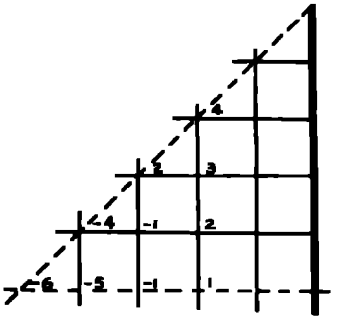


FIGURE 27

5. CURVED BOUNDARIES

When the values of the wanted function are prescribed on a curved boundary, an extra problem is presented in the satisfaction of the difference equation at points adjacent to the boundary This is illustrated in figure 28, where the line of points 4321 cuts the boundary at B, at which point the function is known, but the distance $1B$ is some fraction of the tabular interval h

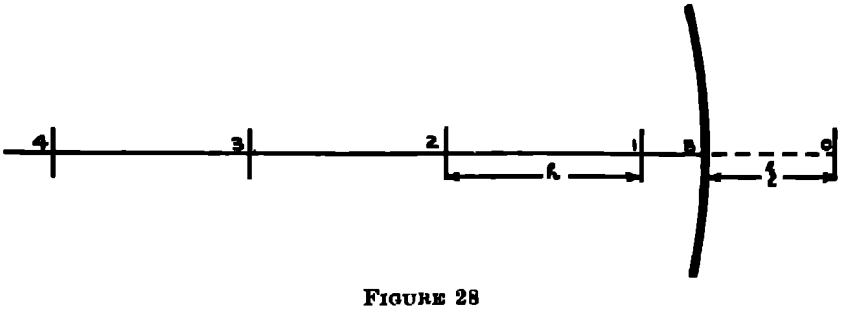


FIGURE 28

It is then required to express a derivative, say the second at the point 1 in terms of the functional values at internal points This is done by producing the line to the point O, where $O1$ is equal to the tabular interval, expressing the value at O in

terms of internal points and the known value at B, and substituting for ϕ_0 in the expression

$$h^2\frac{\partial^2\phi_1}{\partial x^2} = \phi_2+\phi_0-2\phi_1-\tfrac{1}{12}\delta_1^{\text{iv}}+\tfrac{1}{90}\delta_1^{\text{vi}}-\ldots$$

The equation for ϕ_0 is obtained from the Newton-Gregory interpolation formula

$$\phi_B = \phi_0+\binom{x}{1}\Delta_0'+\binom{x}{2}\Delta_0''+\binom{x}{3}\Delta_0'''+\binom{x}{4}\Delta_0^{\text{iv}}+\quad , \tag{23}$$

where

$$\Delta_0' = \phi_1-\phi_0,$$
$$\Delta_0'' = \phi_2-2\phi_1+\phi_0,$$
$$\Delta_0''' = \phi_3-3\phi_2+3\phi_1-\phi_0,$$
$$\text{etc},$$

and $x = \xi/h$ (figure 28)

Neglecting differences higher than the fourth in equation (23), the required expression for ϕ_0 becomes

$$\phi_0 = \frac{\phi_B+\phi_1\left\{-\binom{x}{1}+2\binom{x}{2}-3\binom{x}{3}+4\binom{x}{4}\right\}+\phi_2\left\{-\binom{x}{2}+3\binom{x}{3}-6\binom{x}{4}\right\}+\phi_3\left\{-\binom{x}{3}+4\binom{x}{4}\right\}+\phi_4\left\{-\binom{x}{4}\right\}}{1-\binom{x}{1}+\binom{x}{2}-\binom{x}{3}+\binom{x}{4}}. \tag{24}$$

The resulting finite-difference equations for points near the boundary are thus obtained in terms of internal points only These equations involve more points and are more complicated than the corresponding equations at points remote from the boundary, but the extra difficulty of the relaxation process is small

The method of solution then follows familiar lines

(i) A first approximation $\phi^{(0)}$, taking the specified boundary values, is obtained by neglecting the terms involving differences in the simultaneous equations.

(ii) External points are calculated from equation (24), the function differenced in both directions, and the difference corrections computed

(iii) The difference corrections are inserted as new residuals for the determination of a correction $\phi^{(1)}$, whose boundary values are zero

(iv) The process is repeated until corrections become negligible

Example 7

The method is first illustrated by reworking the problem of example 3, choosing now an interval such that no mesh points fall on the boundary.

In figure 29 the basic interval is 0·4, so that all points adjacent to the boundary are half an interval away from it, that is, the value of x in equation (24) is equal to 0·5 for all external points.

Making use of the symmetry, the following are the six finite-difference equations requiring solution.

$$\left.\begin{aligned} 2B+2A'-4A+409{\cdot}6+\Delta_A &= 0,\\ A+C+E+B'-4B+409{\cdot}6+\Delta_B &= 0,\\ 2B+D+C'-4C+409{\cdot}6+\Delta_C &= 0,\\ 2E+F+C-4D+409{\cdot}6+\Delta_D &= 0,\\ 2B+2D-4E+409{\cdot}6+\Delta_E &= 0,\\ 4D-4F+409\ 6+\Delta_F &= 0, \end{aligned}\right\} \tag{25}$$

where $\Delta_A = -\frac{1}{12}(\delta^{\mathrm{IV}}_{A,x}+\delta^{\mathrm{IV}}_{A,y})+\frac{1}{90}(\delta^{\mathrm{VI}}_{A,x}+\delta^{\mathrm{VI}}_{A,y})-\ldots.$

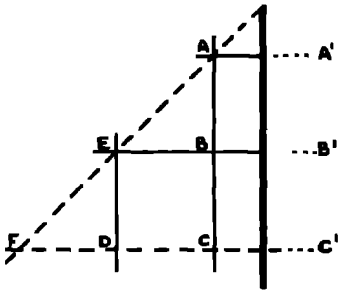


FIGURE 29

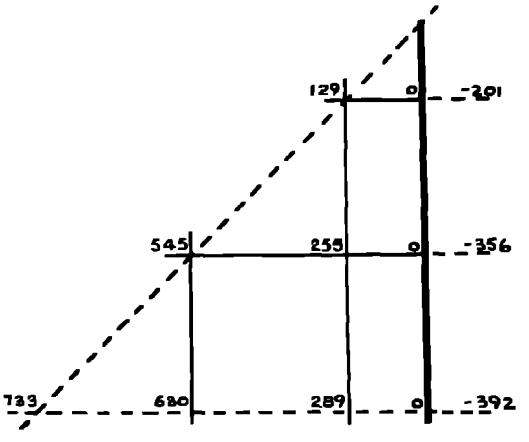


FIGURE 30. $\phi^{(0)}$.

With zero boundary values, equation (24) gives, for the external points A', B', C', the expressions

$$\left.\begin{aligned} A' &= -4A+2{\cdot}1429B-0{\cdot}8C,\\ B' &= -4B+2{\cdot}1429E-0\ 8D,\\ C' &= -4C+2{\cdot}1429D-0{\cdot}8F \end{aligned}\right\} \tag{26}$$

Making these substitutions, the first three of equations (25) becomes

$$\left.\begin{aligned} 6{\cdot}2857B-1{\cdot}6C-12A+409{\cdot}6+\Delta_A &= 0,\\ A+C+3{\cdot}1429E-0{\cdot}8D-8B+409{\cdot}6+\Delta_B &= 0,\\ 2B+3{\cdot}1429D-0{\cdot}8F-8C+409{\cdot}6+\Delta_C &= 0, \end{aligned}\right\} \tag{27}$$

the last three remaining unchanged.

Neglecting the difference correction in equations (27), a first approximation $\phi^{(0)}$ is obtained by relaxation. This solution, together with values at external points computed from equations (26), is shown in figure 30.

When these quantities are differenced (table 12) two fourth-differences are obtained on each line, and the required extrapolation is performed by assuming a constant fifth difference.

TABLE 12

$\phi^{(0)}$	δ'	δ''	δ'''	δ^{iv}	$\phi^{(0)}$	δ'	δ''	δ'''	δ^{iv}	$\phi^{(0)}$	δ'	δ''	δ'''	δ^{iv}
−201					−356					−392				
	+330	—	—	—		+611	—	—	—		+681	—	—	—
+129		−204			+255		−321			+289		−340		
	+126		+112	(−128)		+290		+116	(−92)		+341		+102	(−76)
+255		− 92			+545		−205			+630		−238		
	+ 34		+ 24	− 88		+ 85		+ 35	− 81		+103		+ 32	− 70
+289		− 68		− 48	+630		−170		− 70	+733		−206		− 64

The difference corrections are then calculated and entered in equations (27) as a new residual, from which the correction solution $\phi^{(1)}$ is obtained Figures 31 and 32 show respectively the difference corrections and the solution $\phi^{(1)}$

Further corrections are negligible, and the final solution is given in figure 33, the analytical solution being shown, for comparison, in figure 34

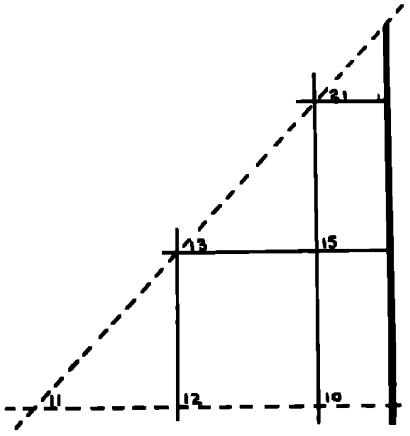


FIGURE 31. $\Delta(\phi^{(0)})$.

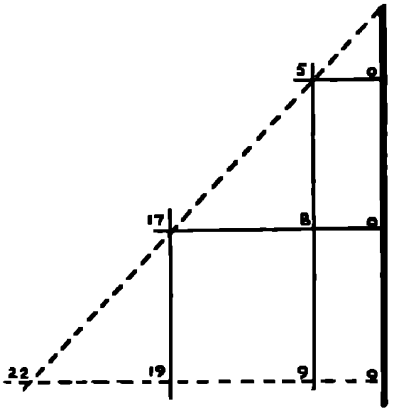


FIGURE 32 $\phi^{(1)}$

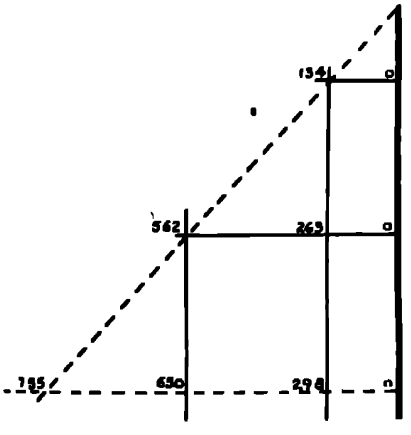


FIGURE 33. Final solution.

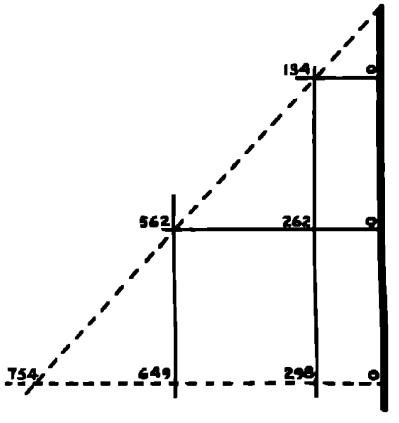


FIGURE 34 Analytical solution.

It should be noted that the differencing of the function not only gives the difference correction, but also enables a rapid estimate to be made, from equation (23), of the accuracy of the formulae used for the satisfaction of the boundary condition. Here,

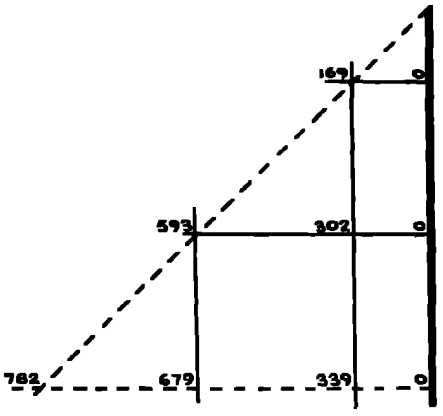


FIGURE 35 (Southwell)

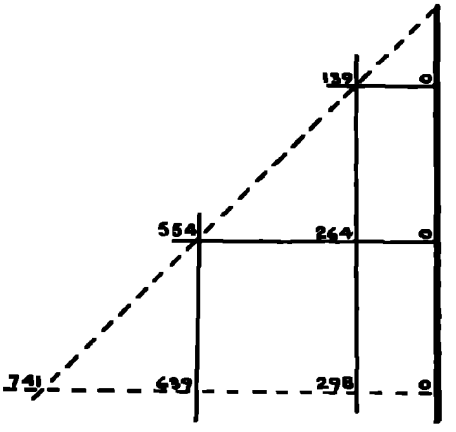


FIGURE 36 (Christopherson)

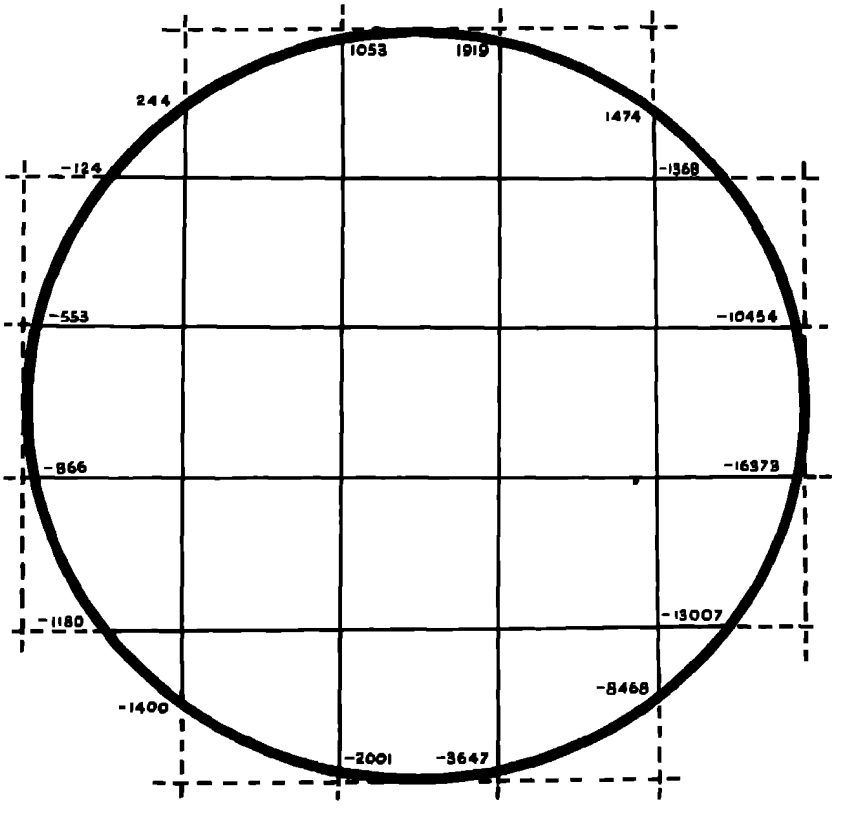


FIGURE 37

for example, the size of the neglected fifth differences indicates that closer accuracy could not be expected on this mesh.

The results obtained by this method may be compared with those given by the methods of Christopherson & Southwell (1938) and Christopherson (1940) The former maintains the simplicity of the method by retaining only the first difference in equation (23), whilst the latter is content to include second differences Both authors ignore the difference correction Figures 35 and 36 give respectively the results given by the methods of Southwell & Christopherson

Christopherson's solution is in this case more accurate than that of Southwell, but it is a fortuitous circumstance that his two sources of error cancel out to a large extent. In the next example, in fact, Christopherson's result is less accurate than Southwell's

Example 8

A sterner test of the method is provided by this last example, in which the differential equation contains a first derivative

It is required to obtain the function which satisfies the differential equation

$$\frac{\partial^2\phi}{\partial r^2}+\frac{1}{r}\frac{\partial\phi}{\partial r}+\frac{\partial^2\phi}{\partial z^2}=0 \tag{28}$$

and takes the values

$$\phi = J_0(3r)\,e^{3z}, \tag{29}$$

on the boundary of a circle with centre at $r = 1\frac{1}{2}$, $z = 0$, and radius 0·5. The analytical solution is of course given everywhere by equation (29)

Figure 37 shows the chosen mesh ($h = 0{\cdot}2$), and the boundary values of the function

In the formulae for the values of external points (equation (24)), there are two independent values of x, and the equations are

$$\left.\begin{aligned} x = 0\ 5, \quad & \phi_0 = 3\ 6571\phi_B - 4\phi_1 + 2\phi_2 - 0{\cdot}8\phi_3 + 0\ 1429\phi_4, \\ x = 0{\cdot}05051, \quad & \phi_0 = 1\ 1130\phi_B - 0\ 2128\phi_1 + 0{\cdot}1556\phi_2 - 0\ 0685\phi_3 + 0{\cdot}0128\phi_4 \end{aligned}\right\} \tag{30}$$

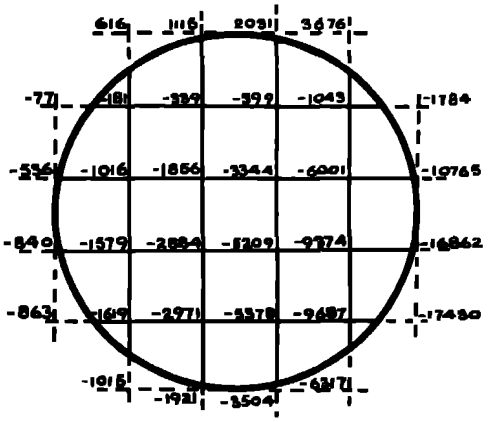


FIGURE 38 $\phi^{(0)}$

Equation (6) is the appropriate finite-difference equation, and the first approximation, obtained by relaxation with the difference correction ignored, is shown in figure 38. Values at external points, computed from equations (30) are included.

When $\phi^{(0)}$ is differenced, two fourth differences are obtainable on each line, and extrapolated values are obtained by assuming the fifth difference constant. Third as well as fourth differences are required in the r-direction. The differences are given in table 13.

TABLE 13

z	$\phi^{(0)}$	δ'	δ''	δ'''	δ^{iv}	δ^{v}	$\phi^{(0)}$	δ'	δ''	δ'''	δ^{iv}	δ^{v}
				$r = 1{\cdot}8$						$r = 1{\cdot}6$		
−0·5	− 77		—	—	—	—	− 536		—	—	—	—
		−104						− 480				
−0·3	− 181		− 54		(−37)		− 1016		− 360		(− 49)	
		−158		− 48		(+3)		− 840		−288		(−184)
−0·1	− 339		−102		−34		− 1856		− 648		−233	
		−260		− 82		+3		−1488		−521		−184
+0·1	− 599		−184		−31		− 3344		−1169		−417	
		−444		−113		(+3)		−2657		−938		(−184)
+0·3	−1043		−297		(−28)		− 6001		−2107		(−601)	
		−741		—		—		−4764		—		—
+0·5	−1784		—		—		−10765		—		—	
				$r = 1{\cdot}4$						$r = 1{\cdot}2$		
−0·5	− 840		—	—	—	—	− 803		—	—	—	—
		− 739						− 756				
−0·3	− 1579		− 566		(− 69)		− 1619		− 590		(− 91)	
		−1305		− 454		(−297)		−1352		− 459		(−297)
−0·1	− 2884		−1020		−306		− 2971		−1055		−388	
		−2325		− 820		−297		−2407		− 847		−297
+0·1	− 5209		−1840		−663		− 5378		−1902		−685	
		−4165		−1483		(−207)		−4309		−1532		(−297)
+0·3	− 9374		−3323		(−960)		− 9087		−3434		(−982)	
		−7488		—		—		−7743		—		—
+0·5	−16862		—		—		−17430		—		—	
r			$z = -0{\cdot}3$						$z = -0{\cdot}1$			
1·0	−1015		—	—	—	—	−1921		—	—	—	—
		−604		(+ 74)				−1050		(+216)		
1·2	−1619		+644		(−195)		−2971		+1137		(−412)	
		+ 40		−121		(+67)		+ 87		−196		(+150)
1·4	−1579		+523		−128		−2884		+ 941		−256	
		+563		−249		+67		+1028		−452		+150
1·6	−1016		+272		− 61		−1856		+ 489		−100	
		+835		−310		(+67)		+1517		−552		(+150)
1·8	− 181		− 38		(+ 6)		− 339		− 63		(+ 50)	
		+797		(−304)		—		+1454		(−496)		—
2·0	+ 616		—		—		+1115		—		—	
			$z = +0{\cdot}1$						$z = +0{\cdot}3$			
1·0	−3504		—	—	—	—	−0317		—	—	—	—
		−1874		(+412)				−3370		(+ 732)		
1·2	−5378		+2043		(−759)		−9087		+3683		(−1355)	
		+ 169		−347		(+290)		+ 313		− 623		(+503)
1·4	−5209		+1696		−469		−9374		+3060		− 852	
		+1865		−816		+200		+3373		−1475		+503
1·6	−3344		+ 880		−179		−6001		+1585		− 349	
		+2745		−995		(+290)		+4958		−1824		(+503)
1·8	− 599		− 115		(+111)		−1043		− 239		(+ 154)	
		+2630		(−884)		—		+4719		(−1670)		—
2·0	+2031		—		—		+3676		—		—	

It is immediately clear from table 13 that the neglect of fifth and higher differences in equation (23), from which equations (30) were derived, will lead to small errors, greatest where the lines $r = 1{\cdot}2$ and $z = 0{\cdot}3$ cut the boundary. The fifth difference on the line $r = 1{\cdot}4$, though of the same magnitude as that on $r = 1{\cdot}2$, is multiplied by a very small coefficient in the interpolation formula, and is therefore not so significant.

The difference corrections are then obtained, using differences up to the fifth in the computation The resulting values are shown in figure 39. The corresponding correction solution $\phi^{(1)}$ is shown in figure 40

Further difference corrections have only a slight effect, and the final solution is shown in figure 41 For comparison, the analytical solution is given in figure 42.

It is seen that the maximum errors occur at points nearest to the boundary, and they are due partly to the extrapolations necessary for the calculation of the difference correction at these points, and partly to the neglect of fifth and higher differences in the formulae used for the satisfaction of the boundary condition

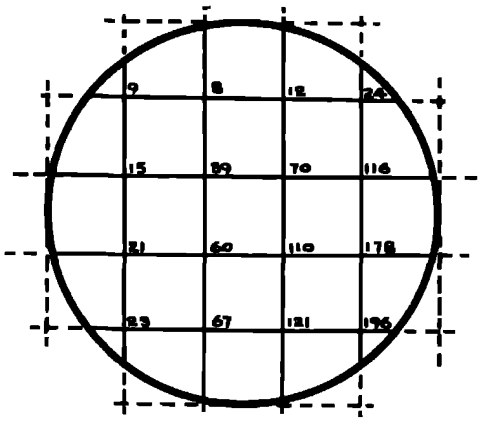


FIGURE 39 $\Delta(\phi^{(0)})$

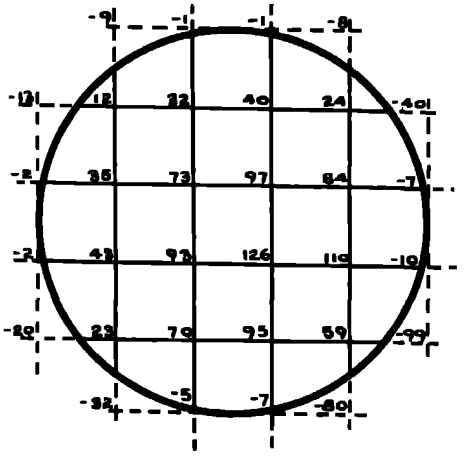


FIGURE 40 $\phi^{(1)}$

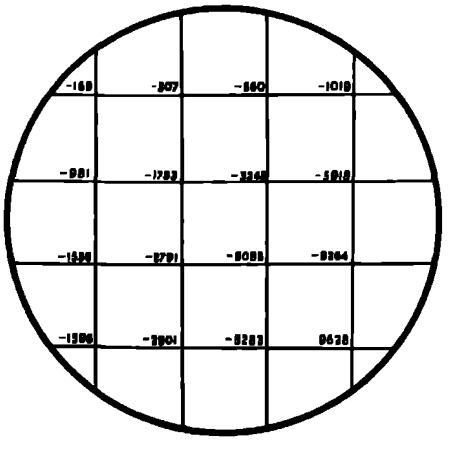


FIGURE 41 Final solution

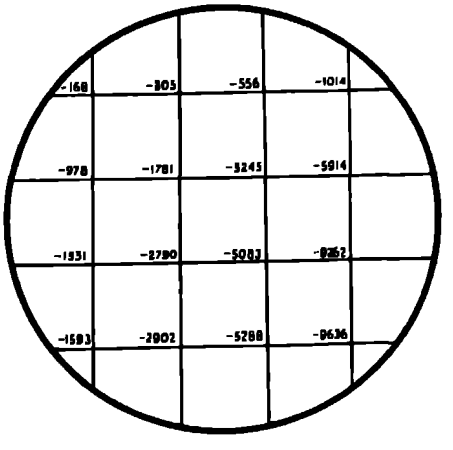


FIGURE 42 Analytical solution

The corresponding results obtained by the methods of Southwell & Christopherson are given in figures 43 and 44 respectively.

The relative accuracy of the three methods is shown in table 14, in which is given the maximum and average arithmetical errors by the three methods. Method 1 is the one used in this paper, method 2 is Southwell's, and method 3, Christopherson's.

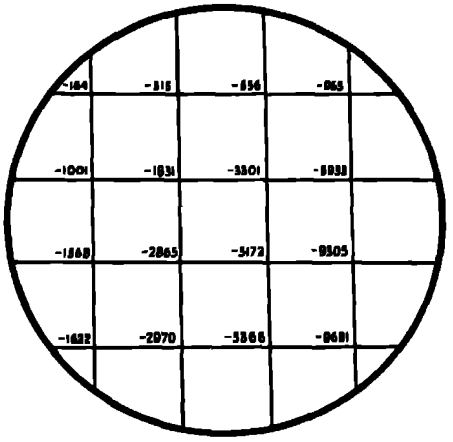


FIGURE 43 (Southwell)

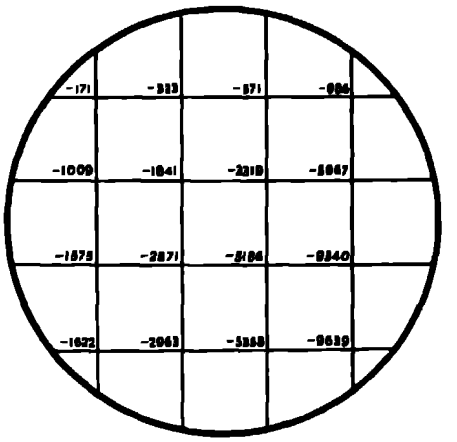


FIGURE 44 (Christopherson)

TABLE 14

	maximum arithmetical error	average arithmetical error
method 1	8	3
method 2	89	40
method 3	103	47

CONCLUSION

In this paper the standard use of relaxation methods has been extended by the inclusion of terms usually neglected in the finite-difference equations involved The solution obtained by neglecting these terms is called the first approximation, from which close estimates for the neglected terms, called the difference correction, are obtained These in turn represent new residuals, from which follow a correction to the first approximate solution By this means the full finite-difference equations are satisfied.

The method is applied to eight examples, covering ordinary and partial differential equations, eigen-value problems, and problems involving curved boundaries. In each case solutions are obtained to a high degree of accuracy The use of coarser meshes, possible by this treatment, greatly reduces the number of equations for solution, and hence better accuracy is achieved with less labour. The use of differences also enables good estimates to be obtained for the position and magnitude of any remaining errors involved in neglecting higher order differences

The work described above has been carried out as part of the research programme of the National Physical Laboratory, and this paper is published by permission of the Director of the Laboratory.

The author desires to acknowledge the suggestions and advice he received from his colleagues, notably Dr A M Turing and Mr M R Hopkins in the Mathematics Division, National Physical Laboratory. He is also particularly indebted to Mr D H. Sadler, from whom he received the basic ideas on the difference correction method.

References

Allen, Fox, Motz & Southwell 1942 *Phil Trans* A, **239**, 488–500.
Allen, Fox & Southwell 1943 *Phil. Trans* A, **239**, 501 537
Christopherson, D. G. 1940 *Trans. Amer. Soc Mech Engrs*, **62**, A1–A4
Christopherson, Fox, Green, Shaw & Southwell 1942 *Phil Trans* A, **239**, 461–487.
Christopherson & Southwell 1938 *Proc. Roy Soc* A, **168**, 317–350
Fox & Southwell 1941 *Phil Trans* A, **239**, 419 460
Gandy & Southwell 1940 *Phil. Trans.* A, **238**, 453 -475.
Michell, J H. 1900 *Proc. Lond. Math. Soc* **31**, 130–146
Shaw & Southwell 1941 *Proc Roy Soc* A, **178**, 1–17.
Southwell, R. V. 1940 *Relaxation methods in engineering science* Oxford University Press
Tables 1942 *Interpolation and allied tables.* H M Stationery Office
Timoshenko, S 1934 *Theory of elasticity* New York McGraw Hill Book Co

The escape of molecules from a plane surface into a still atmosphere

By K J. Brookfield, H. D. N Fitzpatrick, J F Jackson, J B Matthews and E. A Moelwyn-Hughes

(*Communicated by A. R Todd, F.R.S —Received 29 October 1945— Revised 23 January 1946*)

When molecules escape from solids or liquids into a gaseous atmosphere, they do so at a rate which is governed by certain inherent properties of the condensed systems and by the relative values of the saturation pressure of the condensate and the total pressure of the atmosphere. Present considerations are restricted to systems wherein the saturation pressure is a small fraction of the ambient pressure, and to liquids evaporating from plane surfaces A theory is advanced which embraces the diffusion conditions formulated by Maxwell (1890) and the vacuum conditions investigated by Knudsen (1915), and is cast in such a form as to make accessible to experiment the determination of the coefficients of diffusion and accommodation by novel means.

The escape of molecules from the surface of a liquid or solid into a gaseous atmosphere resembles in many respects the dissolution of a solid into a liquid The hypotheses of dissolution advanced by Noyes & Whitney (1897) and by Nernst (1904) have been synthesized into a general theory by Berthoud (1922). A method

closely resembling that of Berthoud has later and independently been applied by Fuchs (1934) to the problem of vaporization from spherical surfaces Following these workers, the mechanism of the escape of molecules from a plane liquid surface may be formulated as follows.

The number of molecules of liquid escaping per second from a plane surface of area S is

$$-\frac{dN}{dt} = Sn_s\bar{\nu} - Sn\bar{u}\alpha, \tag{1}$$

when N = the total number of liquid molecules in the system,

n_s = the number of liquid molecules per unit area of surface,

$\bar{\nu}$ = the average probability per second that a liquid molecule evaporates,

n = the number of molecules per unit volume of vapour,

$u = (kT/2\pi m)^{\frac{1}{2}}$, where T is the absolute temperature, and m is the mass of a single molecule,

and α = the accommodation coefficient

At the equilibrium state, $-dN/dt = 0$ and $n = n^0$, the saturation concentration of vapour, hence

$$n_s\bar{\nu} = n^0\bar{u}\alpha \tag{2}$$

Eliminating $n_s\bar{\nu}$ from equations (1) and (2), then

$$-\frac{dN}{dt} = S\bar{u}\alpha(n^0 - n). \tag{3}$$

Near the surface, a stationary concentration of vapour molecules is rapidly established due to the balancing of the natural escaping tendency of liquid molecules on the one hand, and the hindrance to escape offered by surrounding gas or vapour molecules on the other hand. The stationary concentration, n', of vapour molecules thus established is less than n^0, and, in the stationary state, the rate of vaporization consequently becomes

$$-\frac{dN}{dt} = Su\alpha(n^0 - n'). \tag{4}$$

The rate of loss of molecules due to diffusion, is, by Fick's law,

$$-\frac{dN}{dt} = -SD\left(\frac{\partial n}{\partial x}\right), \tag{5}$$

where D is the coefficient of diffusion, and $\partial n/\partial x$ is the concentration gradient. Application of the general diffusion law

$$\frac{dn}{dt} = D\frac{d^2n}{dx^2}, \tag{6}$$

to the stationary state gives the relation

$$\frac{dn}{dt} = 0 \tag{7}$$

The molecular concentration, n, at a fixed distance, x, from the surface, is thus independent of time. As D itself cannot be zero, it follows that

$$\frac{d^2n}{dx^2} = 0 \tag{8}$$

On integrating, one obtains, for the dependence of the stationary concentration on the distance, x, from the surface, the relation

$$n = ax + b. \tag{9}$$

The constants, a and b, are determined by the boundary conditions

$$n = n', \quad x = \Delta, \tag{10}$$

and

$$n = 0, \quad x = H, \tag{11}$$

so that

$$n = n'\frac{H-x}{H-\Delta}, \tag{12}$$

and consequently

$$\frac{\partial n}{\partial x} = \frac{n'}{H-\Delta}, \tag{13}$$

so that

$$-\frac{dN}{dt} = SD\frac{n'}{H-\Delta}. \tag{14}$$

Combining equations (4) and (14), then

$$n' = n^0\frac{\bar{u}\alpha(H-\Delta)}{\bar{u}\alpha(H-\Delta)+D}. \tag{15}$$

Eliminating n' from equation (15) and either equation (4) or equation (14), the expression

$$-\frac{dN}{dt} = SDn^0[(H-\Delta)+(D/\bar{u}\alpha)]^{-1} \tag{16}$$

is obtained, which reduces, when $(H-\Delta)$ is small, to

$$-\frac{dN}{dt} = S\bar{u}\alpha n^0, \tag{17}$$

as in the theory of Knudsen (1915), and when $(H-\Delta)$ is relatively large ($> 10^{-4}$ cm.) to

$$-\frac{dN}{dt} = SD\frac{n^0}{H-\Delta}, \tag{18}$$

as in the theory of Maxwell (1890). The thickness, Δ, of the diffusion layer must not, of course, be confused with Maxwell's correction factor of $(2/5)d$, which must be added to the height, H, of an open cylinder when it is comparable with the diameter, d.

If w is the weight of liquid vaporizing per second from an area, S, then

$$w = -m\frac{dN}{dt}, \tag{19}$$

and therefore the general expression for the rate of loss of weight becomes

$$w = mSDn^0[(H-\Delta)+D/\bar{u}\alpha]^{-1}, \tag{20}$$

which may conveniently be rewritten as follows:

$$\frac{S}{w} = \frac{(D/\bar{u}\alpha)-\Delta}{mn^0D} + \frac{H}{mn^0D} \tag{21}$$

The measurable quantities (S/w) and H are thus linearly related, and the gradient of the curve should have the constant value of $1/mn^0D$, from which the diffusion coefficient, D, can be obtained, provided the saturation concentration, n^0, is known. Then, in fact,

$$D = \left\{mn^0 \frac{d(S/w)}{dH}\right\}^{-1} \tag{22}$$

The positive intercept made by the curve on the (S/w) axis, corresponding to $H = 0$, should, in principle, allow of an evaluation of the accommodation coefficient, α. In practice, however, it is difficult to measure the height H absolutely, because of the ill-defined nature of the boundary corresponding to $n = 0$. The method, therefore, while yielding reliable values for the diffusion coefficient, D, can yield only approximate values for the accommodation coefficient α.

Apparatus and procedure

A straight-sided glass weighing bottle, A, of cross-sectional area S, is placed inside a wider tube C, with a little mercury in the bottom to ensure thermal contact (see figure 1). A tube B, which slides easily into A, is closed at the bottom with

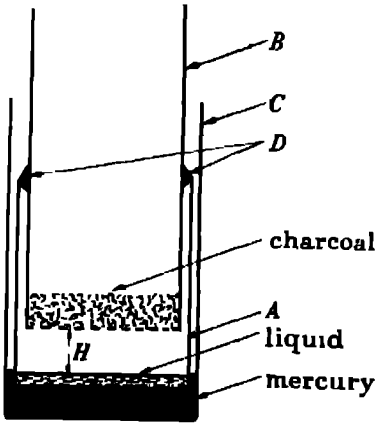


Figure 1. Glass apparatus for measuring diffusion coefficients.

copper gauze, and contains granular, dust-free, freshly degassed charcoal. A known weight of liquid is placed in A, and B is slid down till the gauze is at a height H above the liquid surface, being then arrested by a rim of sealing wax DD. The apparatus

is immersed in a thermostat at a temperature which is maintained constant to within $\pm 0\ 05°$ C. The weighing bottle is weighed and placed in the outer tube. The charcoal tube B is then lowered carefully into position and the time noted After a suitable interval of time the procedure is reversed, the time again being noted, and the loss of weight of the weighing bottle obtained w, the loss of weight per second, is thus found When S/w is plotted against H, a straight line results, in agreement with equation (21) The gradient of the line, obtained by the method of least squares, is equated to $(mn^0D)^{-1}$ Values of n^0 are obtained from the vapour pressure p^0 by assuming the ideal gas laws to hold

$$n^0 = p^0/kT \tag{23}$$

The adopted total pressure is the mean value holding for the period of the runs, and is obtained from the laboratory barometer chart. The height H is determined by cathetometer readings, combined with a knowledge of the density and total amount of liquid employed. The uniform area, S, of the liquid surface is deduced by observing the increase in height occasioned by introducing a known weight of mercury into the vessel

At temperatures above 40° C, the glass apparatus is found to be not so satisfactory because of the softening of the sealing wax, so an improved apparatus was constructed (figure 2) The principle remains the same, but the glass tube containing the charcoal is replaced by a mild steel tube The tube is held in position by means of an adjustable collar, which in one design rested on the top of the glass vessel, and in another on the outer tube (see figure 2) The outer tube is also made of mild steel, and screwed into a heavy base, the joint being made watertight by a rubber washer The height of the charcoal above the liquid is measured by adjusting the collar to one of a series of marks that had been etched on the inner tube, the height from the bottom of the tube having previously been measured with a cathetometer The depth of liquid in the tube is calculated from its weight and density and the area of the glass vessel, and is subtracted from the height corresponding to the position of the etched mark.

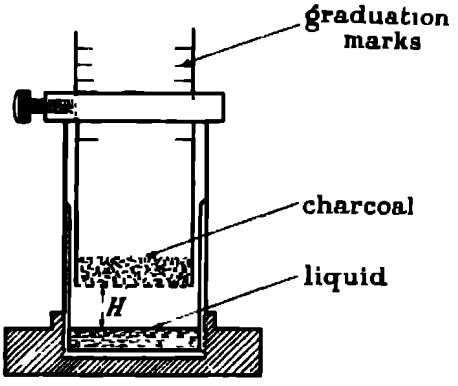


FIGURE 2. Improved mild steel apparatus for measuring diffusion coefficients.

Results

The linear relationship required by equation (21) between (S/w) and H has been established for a number of liquids over a fairly wide range of diffusion conditions. The data given in table 1 have been obtained for two liquids which are typical of a large number examined

Table 1

vapour	t°C	p^0 (mm.)	n^0 (mol/c c)	$d(S/w)/dH$ (cm sec /g)	D(cm²/sec.)
furfural	25 0	2 03	$6\ 61 \times 10^{18}$	$1{\cdot}085 \times 10^{7}$	0·087
	50 0	9·55	$28\ 8 \times 10^{16}$	$0\ 207 \times 10^{7}$	0 107
methyl salicylate	25 0	0 108	$3{\cdot}515 \times 10^{15}$	$1{\cdot}392 \times 10^{7}$	0·082
	50 0	0·717	$21{\cdot}6 \times 10^{15}$	$1\ 304 \times 10^{6}$	0 142

The results are not extensive enough to allow a decision as to the exact form of the temperature dependence of the diffusion coefficient, which, according to the theory of activated diffusion, takes the form

$$D = AT^{3/2}e^{-B/RT}. \tag{24}$$

B, which is known to be small, is therefore assumed to be zero, and the values of A given in table 2 are obtained

Table 2

substance	A	$\sigma_{12} \times 10^{8}$ (cm.)
furfural	$(1{\cdot}77 \pm 0\ 08) \times 10^{5}$	4·54
methyl salicylate	$(2{\cdot}01 \pm 0\ 43) \times 10^{5}$	4·61

According to the Stefan-Maxwell theory, the constant A is given in terms of the sum, σ_{12}, of the molecular radii of the molecule diffusing and the molecules through which it diffuses:

$$A = \frac{2k}{3\pi P\sigma_{12}^2}\left(\frac{2k}{\pi m}\left(1+\frac{m}{m_i}\right)\right)^{\frac{1}{2}} \tag{25}$$

Here, k is Boltzmann's constant, m is the mass of one molecule of liquid, m_i the mass of a 'molecule' of air, and P the total pressure of one atmosphere. The values of σ_{12} derived for furfural and methyl salicylate are given in table 2

Experimental determination of the coefficient of diffusion of water vapour in air

As a check upon the method, the diffusion of water vapour into still air at atmospheric pressure and at various temperatures has been measured Because the diffusing molecules are now lighter than air 'molecules', the apparatus has to be so modified that diffusion takes place downwards. Sulphuric acid replaces charcoal as the absorbing medium. A large, wide-necked bottle, containing water to the

depth of a few centimetres, was placed in a thermostat up to the neck (figure 3) The air inside the bottle was assumed to be always saturated. A straight-sided weighing bottle, with area S, containing concentrated sulphuric acid, the surface of which was at a distance H below the rim, was placed inside the large bottle on a rubber bung which projected above the water surface as an island The dimensions of the large bottle greatly exceeded those of the weighing bottle It was assumed that the concentration of the water vapour was the saturation value at the top of the weighing bottle, and zero at the surface of the sulphuric acid. The weighing bottle was weighed, the stopper removed, and placed in the large bottle After a time t, the weighing bottle was removed, restoppered and weighed again, the gain in weight being noted The ratio S/w, where w is the rate of gain in weight, was plotted against H, and the slope of the straight line found The vapour densities were calculated from the vapour pressures given in Landolt-Bornstein's *Tabellen* (table 3)

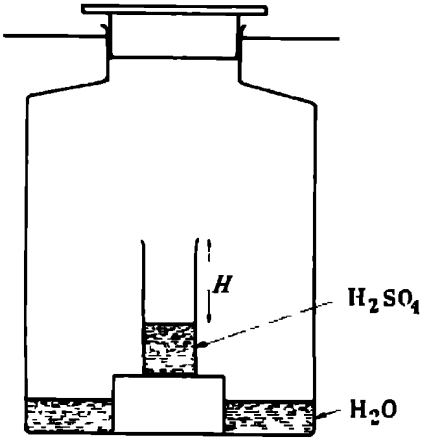


FIGURE 3 Apparatus for measuring diffusion coefficient of water vapour into air

TABLE 3. VAPOUR PRESSURES AND VAPOUR DENSITIES FOR WATER

T(°K)	p^0(mm)	$mn^0 \times 10^5$(g /c c)
298·1	23 76	2·304
308·1	42 19	3·957
318·1	71 90	6 533

The results are summarized in tables 4 and 5 Except at the highest temperature, when the partial vapour pressure is a significant fraction of the total pressure, D is roughly proportional to $T^{\frac{1}{2}}$. The average value of $D/T^{\frac{1}{2}}$ may be taken as $4{\cdot}82 \times 10^{-5}$.

From the Stefan-Maxwell equation

$$D = \frac{2}{3\pi(n+n_i)\sigma_{12}^2}\left\{\frac{2kT}{\pi m}\left(1+\frac{m}{m_i}\right)\right\}^{\frac{1}{2}}, \tag{26}$$

the mean molecular diameter of the water and air 'molecules' is thus found to be $3\ 61 \times 10^{-8}$ cm This value lies between that of 4 14 A found from transport phenomena, and 2·99 A from the constants 'b' of van der Waals's equation.

The data of Winkelmann (1884) and of Trautz & Muller (1935), as summarized by Dorsey (1932), give the figures in table 6 as the most dependable values hitherto obtained for the coefficients of diffusion of water in air.

TABLE 4 DIFFUSION OF WATER VAPOUR INTO AIR

$S = 4{\cdot}5$ cm²

temp. °K	P (mm)	height H cm.	time of run sec.	wt gained g.	w g sec.$^{-1}$	S/w cm^{2}sec g^{-1}	$d(S/w)/dH$ cm sec g.$^{-1}$	D cm.^{2}sec.$^{-1}$
298·1	750	2 1	7200	0 069	$9\ 6 \times 10^{-6}$	$4{\cdot}70 \times 10^{5}$	$1\ 74 \times 10^{5}$	$0{\cdot}250 \pm 0\ 005$
		3 8	9000	0·054	$6{\cdot}0 \times 10^{-6}$	$7\ 53 \times 10^{5}$		
		4 8	13620	0·067	$4{\cdot}92 \times 10^{-6}$	$9\ 18 \times 10^{5}$		
		6 3	6000	0 023	$3\ 8 \times 10^{-6}$	$11\ 9 \times 10^{5}$		
308 1	752·5	1 6	2700	0 058	$2\ 15 \times 10^{-5}$	$2\ 11 \times 10^{5}$	$9\ 54 \times 10^{4}$	$0{\cdot}265 \pm 0\ 005$
		3 55	6300	0 068	$1{\cdot}03 \times 10^{-5}$	$4\ 19 \times 10^{5}$		
		5·1	3600	0·028	$7\ 78 \times 10^{-6}$	$5\ 81 \times 10^{5}$		
		5 55	6300	0 050	$7\ 94 \times 10^{-6}$	$5{\cdot}69 \times 10^{5}$		
		6 5	4800	0·032	$6\ 66 \times 10^{-6}$	$6\ 78 \times 10^{5}$		
318 1	754	1 6	1800	0 068	$3\ 78 \times 10^{-5}$	$1{\cdot}10 \times 10^{5}$	$5\ 37 \times 10^{4}$	$0\ 285 \pm 0\ 005$
		3 4	2700	0 054	$2{\cdot}00 \times 10^{-5}$	$2\ 26 \times 10^{5}$		
		5 15	2700	0·039	$1{\cdot}44 \times 10^{-5}$	$3\ 13 \times 10^{5}$		
		7 2	4500	0 049	$1{\cdot}09 \times 10^{-5}$	$4\ 14 \times 10^{5}$		

TABLE 5 COEFFICIENT OF DIFFUSION OF WATER VAPOUR IN AIR AT 1 ATM.

T(°K)	D(cm^{2}sec.$^{-1}$)	$(D/T^{\frac{3}{2}}) \times 10^{5}$
298·1	0 247	4 80
308·1	0·262	4 84
318 1	0·283	5 00

TABLE 6 COMPARISON WITH PREVIOUS WORK ON WATER

t°C	D(Dorsey)	D(this work)
25	0 257	0 247)
35	0·271	0 262 } ± 0 005
45	0 286	0·283)

The method outlined in this work thus provides reliable values of the coefficient of diffusion, and may yet be developed to yield equally reliable values of the coefficient of accommodation

We are indebted to Professor D Brunt, of the Imperial College of Science and Technology, for valuable criticism of this work, and to the Chief Scientific Officer of the Ministry of Supply for permission to publish it.

References

Berthoud 1922 *J. Chem Phys.* **10**, 633
Dorsey 1932 *The Properties of Ordinary Water*, p 73 Reinhold Publishing Co.
Fuchs 1934 *Phys Z. Sowjet* **6**, 224
Knudsen 1915 *Ann Phys , Lpz* , **47**, 697
Maxwell 1890 *Scientific Papers*, **2**, 635. Cambridge University Press
Nernst 1904 *Z. phys. Chem* **47**, 52.
Noyes & Whitney 1897 *Z phys Chem* **23**, 689
Trautz & Muller 1935 *Ann Phys , Lpz* , **212**, 313
Winkelmann 1884 *Ann Phys , Lpz* , **22**, 152

On the propagation of electromagnetic waves through the atmosphere

By B. K. Banerjea, M Sc , *Palit Laboratory of Physics, University College of Science, Calcutta, India*

(*Communicated by M N Saha, F R S —Received* 18 *July* 1946)

A general method of tackling the problem of the propagation of electromagnetic waves in the ionosphere has been developed and the current methods of Appleton, Hartree, Saha, Rai and Mathur, etc , have been deduced as special cases from the general results The different assumptions by Appleton, Hartree, Bose, Booker and Rai, as regards the condition of reflexion of the waves from the ionosphere, have been shown to be identical A symbol-correspondence chart for the different symbols used by the different workers has been given to facilitate the understanding of the parallelism between the different methods Polarization of the radio waves have been discussed fully

Introduction

The existence of the region of the atmosphere which we now call 'ionosphere' was first postulated by Balfour Stewart (1878) to explain the daily variation of the earth's magnetic field, but it was in 1902 that Kennelly and Heaviside pointed out that the hypothesis of a conducting layer at a height of about 100 km can explain the paradoxical result that wireless signals can, in spite of the curvature of the earth, be propagated over great distances. It was, however, as late as 1925, when its existence was first directly proved by Appleton & Barnett (1925) in England by the slow variation of the transmitter wave-length and, later, by Breit & Tuve (1926) in the U.S.A. by the use of the pulse technique method The latter method, which is now in universal use for investigation of the ionosphere, gives apparently spectacular exhibition of its existence, from the appearance of the reflected echo on the cathode ray oscillograph on a properly calibrated time axis.

The detailed interpretation of these echoes, however, is not so simple, as has been emphasized by various workers, and the reason is twofold. (*a*) experimental, and (*b*) theoretical.

Experimental methods in common use are: (i) The angle of incidence method developed by Appleton & Barnett (1925) to find out the height of the ionosphere consists mainly in comparing the signal intensity received on a vertical with that received on a loop aerial.

(ii) Wave-length change method also developed by Appleton & Barnett (1925). In this method the interference pattern of the ground wave and the sky wave are studied and the number of resultant signal maxima and minima are recorded under an artificial continuous change of phase of the transmitted wave-train.

(iii) Group retardation method Breit & Tuve (1926) developed a novel method of measuring the time taken by the signal group to reach the ionosphere and then to come down to the receiver A pulse of very short duration is transmitted and received at a moderate distance from the transmitter Both the ground wave and the reflected wave are received on the same calibrated time axis of the cathode ray oscillograph. The sharp and regular echo received on the oscillograph screen points towards the existence of a reflecting layer Various other workers like White (1931), Bruewitch (1934), Goubau & Zenneck (1930), Colwell, Friend & Hall (1936), Banerjee (1945) have developed new techniques for intensifying and shortening the duration of the transmitted pulse, and pulses of 20 to 50 kW peak power, and duration of 10^{-5} sec. are at present available for studying the properties of the ionosphere

The appearance of an echo pattern adds weight to the assumption underlying the ray method of tackling the propagation problem that the ionosphere may be regarded as a sharply stratified medium, the width of each stratum being less than the wave used. It may be pointed out that the methods of observation are not yet refined enough to prove the correctness or otherwise of the assumption of gradual stratification in the ionosphere Thus in 'Breit and Tuve experiment' on the group retardation method, let

α mm on the oscillograph screen $\equiv \tau$ sec.

and Δ mm be the least measurable width on the screen $\equiv$ 1 mm.

Then the minimum width of the reflecting layer, as can be inferred from the sharpest echo pattern, is for the best type of transmitting and receiving arrangement $= \tau/\alpha\Delta v \simeq 1$ km $\simeq 10\lambda$ for the E layer, where v = velocity of propagation of the waves Thus even the sharpest echo pattern observed in the oscillograph screen does not justify the assumption underlying the ray method Improved methods of generating very sharp pulses have already been devised, but the received echo is broadened owing to the fact that the band of reception of the receiver itself has a finite width. Still with these specially designed sharp pulse transmitters, Colwell, Friend & Hall (1936) have observed diffused echoes, showing that the wave is reflected not from a sharp layer, but from a diffused layer.

The theoretical difficulties are of a fundamental nature, for though after the great success of the group retardation method for investigating the properties of the ionosphere by sending wireless waves to and receiving them back from the ionosphere both vertically and obliquely, the actual problem of the propagation of the electro-

magnetic waves through the ionosphere has not been solved rigorously The reasons are

(*a*) The equation of propagation of electromagnetic waves through the ionosphere had not been rigorously developed

(*b*) The boundary conditions, by which are meant the arrangement of the antenna, the conditions of the ground on which the antenna is situated and the physical conditions of the ionosphere through which the electromagnetic waves are propagated, are difficult to be properly formulated and utilized in the solutions of the equations

The theoretical method most in vogue is the magneto-ionic theory developed by Appleton (1925) in England. It is also known as the ray method The final form to this theory has been given by Appleton (1932) and Hartree (1931) and discussed further by Appleton (1935) A critical review of the methods based on the magneto-ionic theory is given below.

Present theoretical methods

In order to compare the merits of the different methods as well as to show their interconnexion, the different works will be transcribed to the same uniform notation and the different reference framework will be shown in the same diagram In fact, it has been shown that the works of Appleton (1932), Hartree (1931), Saha, Rai & Mathur (1937) and Saha & Banerjea (1945) may be deduced from the same general results At a later stage, a symbol correspondence chart will be given which should enable the reader to follow the works of different authors in their original form

The theoretical work done on this subject mainly consists of two parts, viz

(1) Solution of the equation of motion of the charged particles in the earth's magnetic field, under the influence of the electric vector of the incident electromagnetic wave This displacement $\rho(\xi, \eta, \zeta)$ has been used by the different workers for different purposes

(2) Solution of the Maxwell equations

1. *Solution of the equation of motion*

Taking a general framework XYZ the equation of motion of an electron in the earth's magnetic field H under the influence of the incident electric vector $Ee^{i\nu t}$ is

$$\ddot{\rho} + \nu\dot{\rho} - \frac{e}{m_0 c}[\dot{\rho} \times \mathrm{H}] = \frac{e}{m_0} E e^{i\nu t}, \tag{1}$$

where $\rho(\xi, \eta, \zeta)$ is the displacement of the electron, ν the collisional frequency, e, m_0 the charge and mass of the electron (the sign of e has been already taken into account in the above equation) and c the velocity of light The effect of Lorentz' polarization term has been neglected, as the discussions of Darwin (1934) have shown it to be negligible

Assuming that $\rho \propto e^{ipt}$, equation (1) becomes

$$\rho(-p^2+i\nu p)-\frac{iep}{m_0 c}[\rho \times \mathbf{H}] = \frac{e}{m_0} E e^{ipt}$$

Introducing a new dyad $[\mathbf{I} \times \mathbf{H}]$ given by the matrix

$$[\mathbf{I} \times \mathbf{H}] = \begin{vmatrix} 0 & -H_z & H_y \\ H_z & 0 & -H_x \\ -H_y & H_x & 0 \end{vmatrix}, \qquad \mathbf{I} = \begin{vmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{vmatrix},$$

we can express the vector product $[\rho \times \mathbf{H}]$ as the scalar product of the dyad $[\mathbf{I} \times \mathbf{H}]$ with ρ, i e

$$-[\rho \times \mathbf{H}] = [\mathbf{I} \times \mathbf{H}]\ \rho$$

Then the equation of motion becomes

$$-\frac{m_0 p^2}{e}\left(\beta \mathbf{I} - \frac{ie}{m_0 pc}[\mathbf{I} \times \mathbf{H}]\right)\rho = E e^{ipt}, \tag{2}$$

where $\beta = 1 - i\nu/p$.

Putting $\omega = \dfrac{e\mathbf{H}}{m_0 cp}$ and breaking the above vector equation into components, we get

$$\begin{vmatrix} E_x \\ E_y \\ E_z \end{vmatrix} e^{ipt} = -\frac{m_0 p^2}{e} \begin{vmatrix} \beta & i\omega_z & -i\omega_y \\ -i\omega_z & \beta & i\omega_x \\ i\omega_y & -i\omega_x & \beta \end{vmatrix} \begin{vmatrix} \xi \\ \eta \\ \zeta \end{vmatrix}. \tag{3}$$

We can also express ρ in terms of E either by finding out the dyad reciprocal to $\beta \mathbf{I} - \dfrac{ie}{m_0 cp}[\mathbf{I} \times \mathbf{H}]$ by the usual method or from (3) by the usual determinant method. Thus

$$\begin{vmatrix} \xi \\ \eta \\ \zeta \end{vmatrix} = \frac{\mathbf{A}}{Ne} \begin{vmatrix} \beta^2-\omega_x^2 & \omega_x\omega_y+i\beta\omega_z & \omega_x\omega_z-i\beta\omega_y \\ \omega_y\omega_x-i\beta\omega_z & \beta^2-\omega_y^2 & \omega_y\omega_z+i\beta\omega_x \\ \omega_z\omega_x+i\beta\omega_y & \omega_z\omega_y-i\beta\omega_x & \beta^2-\omega_z^2 \end{vmatrix} \begin{vmatrix} E_x \\ E_y \\ E_z \end{vmatrix} e^{ipt} \tag{4}$$

or $\rho = \mathbf{S}.E$, where S is the dyad reciprocal to $-\dfrac{m_0 p^2}{e}\left(\beta \mathbf{I} - \dfrac{ie}{m_0 cp}[\mathbf{I} \times \mathbf{H}]\right)$ occurring in equation (2) and given by the matrix in (4) with

$$\mathbf{A} = \frac{r}{4\pi\beta}\frac{1}{\beta^2-\omega^2}, \quad r = \frac{p_0^2}{p^2} = \frac{4\pi N e^2}{m_0 p^2},$$

N = number of ions per c c.

From the matrix for S, we can write

$$\mathbf{S} = -\frac{e\beta}{m_0 p^2(\beta^2-\omega^2)}\left(\mathbf{I} + \frac{ie}{m_0 cp\beta}[\mathbf{I} \times \mathbf{H}] - \frac{e^2}{m_0^2 c^2 p^2 \beta^2}(\mathbf{HH})\right), \tag{5}$$

where (HH) is the dyadic product of the vector H Instead of using the components of the displacement vector ρ we can use the polarization vector P defined by

$$P = 4\pi Ne\rho = 4\pi Ne\mathrm{S}.E = \sigma_h.E, \tag{6}$$

where σ_h is called 'the scattering tensor' by Darwin.

Different workers used different frameworks and as a result, the components of the vector P are expressed in different forms XYZ (figure 1) is any general framework with OZ as vertical, this framework has been used by Hartree (1932) and Saha, Rai & Mathur (1937) OH is the direction of the magnetic field and $ZX'OH$ is the magnetic meridian with OX' as the horizontal projection of the magnetic line of force. OY' is perpendicular to the magnetic meridian. Let OZ' be the line in $ZX'OH$, the magnetic meridian, perpendicular to OH Appleton (1932) used the framework $X'Y'Z$ in which, for propagation along OZ, $H_{x'}$ = component of the magnetic field transverse to the direction of propagation = H_T, $H_{y'} = 0$, H_z = component along the direction of propagation Saha & Banerjea (1945) used the principal axes $HY'Z'$ of the tensor S in which $\omega_H = \omega_{x'}$, $\omega_{y'} = \omega_{z'} = 0$ and $\angle ZOZ' = \delta$, the dip angle.

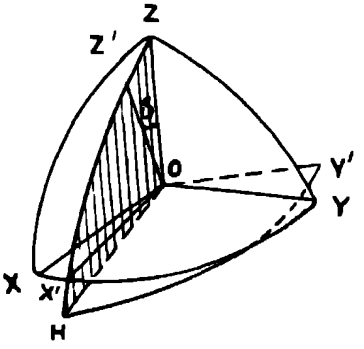


FIGURE 1

The components of polarization vector P were obtained explicitly in this form (4) by Saha, Rai & Mathur (1937). Hartree (1931) obtained the expression for the displacement vector in which the electrons are assumed to be bound by a quasi-elastic force proportional to k_1^2, as the ions in the ionosphere are free, we put $k_1^2 = 0$ and then Hartree's expression for the displacement vector is identical with (4). Appleton's results are obtained from (3) by putting $\omega_{x'} = \omega_T$, $\omega_{y'} = 0$, $\omega_L = \omega_L$

$$\left.\begin{aligned} E_{x'} &= -\frac{1}{\gamma}[\beta P_{x'} + \nu\omega_L P_{y'}], \\ E_{y'} &= -\frac{1}{\gamma}[-i\omega_L P_{x'} + \beta P_{y'} + \nu\omega_T P_z], \\ E_z &= -\frac{1}{\gamma}[-i\omega_T P_{y'} + \beta P_z], \end{aligned}\right\} \tag{6'}$$

Saha & Banerjea's results also follow from (3) by putting $\omega_H = \omega$, $\omega_{y'} = \omega_{z'} = 0$:

$$E_H = -\frac{\beta}{\gamma} P_H, \quad (E_{y'} \pm \imath E_{z'}) = -\frac{\beta \pm \omega}{\gamma}(P_{y'} \pm \imath P_{z'}). \tag{7}$$

2 *Solution of the Maxwell equation*

Next we proceed to set up the propagation equations for the electric and magnetic vectors of the electromagnetic wave from the Maxwell's equations, viz.

$$\nabla \times H = \frac{1}{c}\dot{D}, \qquad \nabla \times E = -\frac{1}{c}\dot{H},$$

$$\nabla_0 H = \nabla_0 D = 0, \qquad D = E + P. \tag{8}$$

Then

$$\nabla \times \nabla \times E + \frac{1}{c^2} \quad = 0, \tag{9}$$

and

$$\nabla \times \nabla \times H - \frac{1}{c}\nabla \times \dot{D} = 0 \tag{10}$$

Also from (6), since $P = 4\pi N e \mathsf{S} E$, we get

$$D = E + 4\pi N e \mathsf{S} E = (\mathsf{I} + 4\pi N e \mathsf{S}) \; E = \mathsf{K}.E, \tag{11}$$

where K is the complex dielectric tensor and given by the matrix.

$$\mathsf{K} = \begin{vmatrix} 1 - 4\pi A(\beta^2 - \omega_x^2) & 4\pi A(\omega_x\omega_y + \imath\beta\omega_z) & 4\pi A(\omega_x\omega_z - \imath\beta\omega_y) \\ 4\pi A(\omega_x\omega_y - \imath\beta\omega_z) & 1 - 4\pi A(\beta^2 - \omega_y^2) & 4\pi A(\omega_z\omega_y + \imath\beta\omega_x) \\ 4\pi A(\omega_x\omega_z + \imath\beta\omega_y) & 4\pi A(\omega_z\omega_y - \imath\beta\omega_x) & 1 - 4\pi A(\beta^2 - \omega_z^2) \end{vmatrix} \tag{12}$$

$$= \begin{vmatrix} 1 - \frac{\gamma}{\beta} & 0 & 0 \\ 0 & 1 - \frac{\gamma\beta}{\beta^2 - \omega^2} & \frac{\imath\gamma\omega}{\beta^2 - \omega^2} \\ 0 & -\frac{\imath\gamma\omega}{\beta^2 - \omega^2} & 1 - \frac{\gamma\beta}{\beta^2 - \omega^2} \end{vmatrix} \tag{12'}$$

the former form (12) being with respect to the general co-ordinates, while the latter (12′) with respect to the principal axes OH, OY', OZ', as used by Saha & Banerjea (1945).

The vectors E, H and D occurring in the Maxwell equations (9) and (10) are the field vectors within the ionosphere The total electric field E, used in (9), however, may be looked upon, as suggested by Darwin (1924), as the superimposition of the incident field E_i and the resultant scattered field E_s

Let us next consider incident wave given by the electric vector

$$E_i = E_0 e^{\imath(pt - k\mathbf{N}_0\mathbf{r})},$$

where $\mathbf{N}_0$ is the wave-normal vector with components l, m, n, $\mathbf{r}$ the propagation vector, and k is the wave number p/c. If we picture the ionosphere as a homogeneous

refracting medium, separated from the region where $\mu = 1$, by a sharp boundary (or stratified layer), and consider the wave travelling from below without change of wave form, we can put the refracted wave as represented by

$$\left.\begin{aligned} \boldsymbol{E} &= \boldsymbol{E}_0 e^{i(pt - k\mathbf{N}_0\mathbf{r})}, \\ \boldsymbol{D} &= \boldsymbol{D}_0 e^{i(pt - k\mathbf{N}_0\mathbf{r})}, \\ \boldsymbol{H} &= \boldsymbol{H}_0 e^{i(pt - k\mathbf{N}_0\mathbf{r})}, \end{aligned}\right\} \tag{13}$$

where $\mathbf{N}$ is a vector with components l, m and q For vertical propagation $\mathbf{N}_0\mathbf{r} = qz$ and q becomes equal to the complex refractive index

Then, from (13), $$[\nabla \times \quad] = -ik[\mathbf{N} \times \quad],$$

$$[\nabla \times \nabla \times \quad] = k^2[\mathbf{N}^2 \mathsf{I} - (\mathbf{NN})].$$

Thus the wave equation for the electric vector becomes

$$[\mathbf{N}^2 \mathsf{I} - (\mathbf{NN})]\, \boldsymbol{E} = \boldsymbol{D}, \tag{14a}$$

or $$[(\mathbf{N}^2 - 1)\, \mathsf{I} - (\mathbf{NN})]\, \boldsymbol{E} = \boldsymbol{P}, \tag{14b}$$

or $$[\mathsf{K} - \mathbf{N}^2 \mathsf{I} + (\mathbf{NN})]\, \boldsymbol{E} = 0 \tag{14c}$$

Similarly for the magnetic vector

$$[(\mathbf{N}^2 - 1)\, \mathsf{I} - (\mathbf{NN})]\boldsymbol{H} = [\mathsf{I} \times \mathbf{N}]\, \boldsymbol{P}. \tag{15}$$

Equations (14) and (15) are the most general equations of propagation of the electromagnetic field vectors in a homogeneous ionosphere Equations used by Hartree (1931), Appleton (1932), Saha, Rai & Mathur (1937), etc , can be obtained as special cases of (14) and (15).

Thus Hartree confines himself to a wave polarized in the XZ plane and inclined at an angle θ with OZ Then the components of the vector $\mathbf{N}$ occurring in (14) and (15) becomes $\sin\theta$, 0, q and (14a) becomes

$$\left.\begin{aligned} D_x &= q^2 E_x - q \sin\theta E_z, \\ D_y &= (q^2 + \sin^2\theta)\, E_y, \\ D_z &= -q \sin\theta E_x + \sin^2\theta E_z, \end{aligned}\right\} \tag{16}$$

which are identical with equations of Hartree's (1931) work Hartree did not work out the corresponding equation for the magnetic vector.

Appleton considered the case of vertical propagation of a linearly polarized wave and then the components of $\mathbf{N}$ are O, O, q, where q is the complex refractive index. Then (14b) becomes,

$$\left.\begin{aligned} (q^2 - 1)\, E_x &= \quad P_x, \\ (q^2 - 1)\, E_y &= \quad P_y, \\ E_z &= -P_z, \end{aligned}\right\} \tag{17}$$

and (15) becomes

$$\left.\begin{aligned}(q^2-1)H_x &= qP_y,\\ (q^2-1)H_y &= -qP_x,\\ H_z &= 0.\end{aligned}\right\} \tag{18}$$

These equations are identical with equations (26) of Appleton (1932) Equations (17), again, can be obtained from (16) by putting $\theta = 0$ when Hartree's case coincides with Appleton's.

In the Saha-Banerjea framework (1945) for vertical propagation, along OZ the components of N along OH, OY' and OZ' are $-q\sin\delta$, O, $q\cos\delta$, where δ is the dip angle Correspondingly, the wave equations become

$$\left.\begin{aligned}D_H &= q^2\cos^2\delta E_H - q^2\sin\delta\cos\delta E_{z'},\\ D_{y'} &= q^2 E_{y'},\\ D_{z'} &= -q^2\sin\delta\cos\delta E_H + q^2\sin^2\delta E_{z'},\end{aligned}\right\} \tag{19}$$

and

$$\left.\begin{aligned}(q^2-1)H_H &= q\cos\delta P_{y'},\\ (q^2-1)H_{y'} &= -q\cos\delta P_H - q\sin\delta P_{z'},\\ (q^2-1)H_{z'} &= -q\sin\delta P_{y'}\end{aligned}\right\} \tag{20}$$

We can also evaluate the expression for the complex refractive index from equation (14c), which states that the determinant of the tensor $[K - N^2 I + (NN)]$ must vanish. Since the value of refractive index does not depend on the direction of propagation of the wave and since the value of a determinant is independent of the axes used, we confine ourselves to vertical propagation and to Saha-Banerjea framework, which is the principal framework and so four out of nine components of K vanish Thus $N_H = -q\sin\delta$, $N_{y'} = 0$, $N_{z'} = q\cos\delta$, where q is the required complex refractive index. The refractive index equation becomes

$$\begin{vmatrix} 1-\frac{r}{\beta}-q^2+q^2\sin^2\delta & 0 & -q^2\sin\delta\cos\delta \\ 0 & 1-\frac{r\beta}{\beta^2-\omega^2}-q^2 & \frac{ir\omega}{\beta^2-\omega^2} \\ -q^2\sin\delta\cos\delta & -\frac{ir\omega}{\beta^2-\omega^2} & 1-\frac{r\beta}{\beta^2-\omega^2}-q^2+q^2\cos^2\delta \end{vmatrix} = 0$$

or

$$q^2 = 1 - \frac{2r}{2\beta + \frac{\omega^2\cos^2\delta}{(r-\beta)} \pm \sqrt{\left[\frac{\omega^4\cos^4\delta}{(r-\beta)^2} + 4\omega^2\sin^2\delta\right]}}, \tag{21}$$

which is identical with Appleton's equation.

Polarization of the radio waves

The electromagnetic waves in the ionosphere are not strictly transverse with respect to the electric vector for $(E\ \mathbf{r}) \neq 0$ but they are transverse with respect to the magnetic vector as $(H\ \mathbf{r}) = 0$ Thus, it has become customary to express the nature of the polarization of the radio waves in terms of the magnetic vector From (13) and (14) it can be easily shown that

$$f = \frac{H_x}{H_y} = \frac{i}{\omega_L}\left(\frac{r}{1-q^2} - \beta\right) = \rho e^{-i\phi}, \quad \text{say}$$

Eliminating $(1-q^2)$ from this and equation (14) we get,

$$f + \frac{1}{f} = \frac{i}{\omega_L} \frac{\omega_T^2}{r} - \beta$$

This is a quadratic equation in f, and, if f_1, f_2 be the roots,

$$f_1 f_2 = 1, \quad \text{i.e.} \quad \rho_1 = 1/\rho_2, \quad \phi_1 = -\phi_2 \tag{22}$$

In order to get the actual form of the polarization ellipse, we put H_x = real part of Ae^{ipt} and H_y = real part of $\rho A e^{i(pt-\phi)}$ Eliminating the time factor pt we get

$$H_x^2 - \frac{2H_x H_y \cos\phi}{\rho} + \frac{H_y^2}{\rho^2} = A^2 \sin^2\phi \tag{23}$$

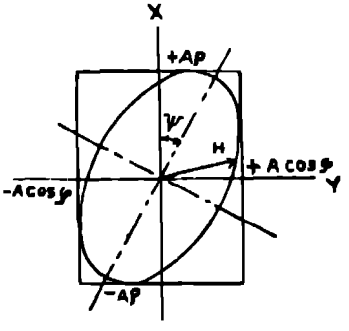


Figure 2

Since the discriminant of this quadratic form is negative, the vector whose components are H_x, H_y is an ellipse in the XY plane. The points of contact of the ellipse with the circumscribed rectangle are $(\pm A\rho, \pm A\cos\phi)$, $(\pm A\cos\phi, \pm A\rho)$ In general the axes of the ellipse do not coincide with the co-ordinate axes but the two systems may be brought to coincidence by rotating the ellipse through an angle ψ (called the tilt of the polarization ellipse) about the Z-axis, such that the product of the components of the vector H in the new co-ordinates vanish. This gives

$$\tan 2\psi = -\frac{2\rho\cos\phi}{1-\rho^2}.$$

The second ellipse corresponding to the other root of the equation (22) is given by

$$\frac{H_x^2}{\rho^2} - \frac{2H_xH_y\cos\phi}{\rho} + H_y^2 = \frac{A}{\rho^2}\sin^2\phi. \tag{24}$$

Since $H_x = Re\, Ae^{ipt}$, $H_y = Re\, Ae^{i(pt-\phi)}$, for ϕ positive the sense of rotation is clockwise and for ϕ negative the sense is anticlockwise

Let us next find out the values of ρ and ϕ. From the relation

$$\frac{i}{\omega_L}\frac{\omega_T^2}{r-\beta} = \rho e^{-i\phi} + \frac{1}{\rho}e^{i\phi}$$

we can obtain, by equating the real and imaginary parts,

$$(r-1)(\rho+1/\rho)\cos\phi + (\nu/p)(\rho-1/\rho)\sin\phi = 0,$$

$$(\nu/p)(\rho+1/\rho)\cos\phi - (r-1)(\rho-1/\rho)\sin\phi = \omega_T^2/\omega_L^2.$$

From these two expressions it is possible to find out the values of ρ and ϕ, but the equations become somewhat unwieldy Let us first take the case where friction ν can be neglected as this case leads to considerable simplifications. We have from above

$$\phi = \tfrac{1}{2}\pi,$$

$$\rho = \frac{\omega\cos^2\delta}{2(1-r)\sin\delta}\left\{1 + \sqrt{\left(1 + \frac{4\sin^2\delta(1-r)^2}{\omega^2\cos^4\delta}\right)}\right\}. \tag{25}$$

This expression was obtained by Saha, Rai & Mathur (1937) and was discussed fully by them, with special reference to the polarization of waves of 100 m length at selected stations ranging from the magnetic north to the magnetic south pole

These relations have been verified by Appleton & Ratcliffe (1928) in England They put $r = 0$ for the reflected wave and get

$$\rho_1 = 0\ 9633, \quad \rho_2 = -1{\cdot}0373 \quad \text{for} \quad \lambda = 100\,\text{m}$$

Both the O-wave and the x-wave are therefore nearly circularly polarized and the sense of rotation for the O-wave is anti-clockwise showing ϕ negative as we would expect from (24)

For Australia, δ_L = negative, we have $\rho_1 = -0\ 9732$, $\rho_2 = 1\ 0274$, both the O- and the x-waves are again circularly polarized as observed by Green (1932)

For equatorial region, there is only one observation, due to Berkner & Wells (1937) at Huancayo They found $\rho_1 = 1/0\ 129$, $\rho_2 = 1/7{\cdot}643$ with the O-wave linearly polarized along the magnetic north-south and the X-wave along the magnetic east-west

In order to link up the different methods the following transcription table is helpful.

UNIFICATION OF NOTATION OF THE DIFFERENT WORKERS

quantity	Appleton and his school	Hartree	Saha and his school
field vectors	$\boldsymbol{D}, \boldsymbol{E}, \boldsymbol{H}$ P	$\boldsymbol{D}, \boldsymbol{L}, \boldsymbol{H}$	$\boldsymbol{D}, \boldsymbol{E}, \boldsymbol{H}$, P
earth's field	H	H	H
refractive index	μ	μ	μ
pulsatance	p	Kc	p
gyrofrequency, $\frac{e\mathrm{H}}{m_0 c}$	$\mathrm{p_h}$	$K_h c$	$\mathrm{p_h}$
collision frequency	ν	$2K_r c$	ν
displacement of the ion	(x, y, z)	ρ	(ξ, η, ζ)
dipole moment	$Ne(x, y, z)$	$Ne\rho$	$Ne(\xi, \eta, \zeta)$
complex refractive index	cq	K	q
conductivity	$\sigma_\parallel, \sigma_\perp$		σ
phase velocity	v	—	v
group velocity	w	—	w
Abbreviations used			
$\frac{4\pi Ne^2}{m_0}$	p_0^2	—	p_0^2
Lorentz factor	α	β	0
$\frac{m_0 p^2}{4\pi Ne^2} = \frac{p^2}{p_0^2}$	$-\alpha$	$\frac{1}{\sigma_0} - \beta$	$\frac{1}{r}$
$\frac{m_0 p}{4\pi Ne^2}\left[\frac{e\mathrm{H}}{m_0 c}\right] = \frac{p p_h}{p_0^4}$	γ	$\frac{\rho}{\sigma_0}$	$\frac{\omega}{r}$
$\frac{e\mathrm{H}}{m_0 c(p - i\nu)}$	—	$k\alpha\mathrm{H}$	$-\frac{\omega}{\beta}$
$\frac{e\mathrm{H}}{m_0 cp} = \frac{p_h}{p}$	$-\frac{\gamma}{\alpha}$	τ	ω
$\frac{4\pi Ne^2}{m_0 p^2(1 - i\nu/p)} \Big/ \left[1 - \frac{p_h^2}{p^2(1 - i\nu/p)^2}\right]$,	$\frac{\alpha}{\alpha^2 - \gamma^2}$	ξ	$-\frac{4\pi A}{\beta^2}$
$1 - i\nu/p$	—	—	β
$\frac{4\pi Ne^2}{m_0 p^2(1 - i\nu/p)}$	$\frac{1}{\alpha + \alpha}$	σ_0	$-\frac{r}{\beta}$
$\frac{e}{m_0 c^2(-p^2 + i\nu p)}$	—	α	$-\frac{e}{m_0 p^2 \beta}$
scattering tensor	—	σ	$\frac{r\beta A}{\beta^2 - \omega^2}$
$\frac{m_0 p\nu}{4\pi Ne^2}$	β	—	$\frac{\delta}{r}$

Conditions of reflexion of the waves

In order to explain the return of the radio waves from the ionosphere different workers have used different criteria.

(i) Appleton & Hartree take that the wave gets reflected when the refractive index vanishes and obtain the well-known relation $p^2 = p^2$ and $p^{-2} = p_0^2 \pm pp_A$ as the conditions for reflexions.

(ii) Booker (1936) assumes that the wave is reflected when the Poynting vector S becomes perpendicular to the wave normal N, i e for vertical propagation, the wave is reflected when the energy vector becomes horizontal.

(iii) In order to explain the fourth condition of reflexion, observed by Bajpai & Pant (1937) at Allahabad, Rai (1937) assumes that the vanishing of the group velocity should be regarded as the proper criterion for reflexion He obtained in addition to Appleton-Hartree conditions, a fourth relation $p_0^2 = p^2 \dfrac{p^2 - p_H^2}{p^2 - p_L^2} \simeq p^2 - p_T^2$ (for $p_L \ll p_T$) which explains the results of Bajpai & Pant

(iv) Bose (1938) suggested that the wave gets reflected when the magnetic vector H is zero. This can take place, either when E is parallel to N or $v/c \simeq \infty$ The first condition is, as will be shown presently, equivalent to Appleton's condition

All the above criteria are *assumptions*, and though probably they give us a rough approximation to the actual process of wave propagation in the ionosphere, they cannot be regarded as strictly mathematical solutions to the Maxwellian equations, not even for a homogeneous absorbing ionosphere Further, in an absorbing medium as the ionosphere is, μ can never become zero.

It is, however, shown below that the above four assumptions, though apparently different and probably intended by their authors to represent different criteria, are actually identical, when friction is neglected

From the Maxwell's field equations, for a homogeneous ionosphere

$$\partial H/\partial t = -c\nabla \times [E_0 e^{ip(t - r\mathrm{N}/v)}],$$

where N wave normal, r propagation vector. Therefore

$$H = H_0 e^{ip(t - r\mathrm{N}/v)}, \quad H_0 = ipc/v[\mathrm{N} \times E_0].$$

Thus the magnetic vector is normal to the plane defined by N and E Again

$$\partial D/\partial t = c\nabla . H, \quad \text{i e} \quad D = D_0 e^{ip(t - r\mathrm{N}/v)},$$

where $$D_0 = [H \times \mathrm{N}] = c/v\,[-\mathrm{N}(\mathrm{N}\,E_0) + E_0].$$

Thus D is normal to the plane defined by H_0, and N and is coplanar with the vectors N and E. Hence, as in crystals, the wave is transverse in D and H The vector E makes an angle Ψ with D such that

$$\cos \Psi = \frac{v}{c}\frac{|D|}{|E|} = \frac{v}{c}\left[1 - \frac{4\pi\sigma}{p}\right].$$

The Poynting vector $S = c/4\pi[E \times H]$ is evidently coplanar with D, E and N and perpendicular to E, i e makes the same angle Ψ with N The actual orientations of the field vectors are as shown in figure 3

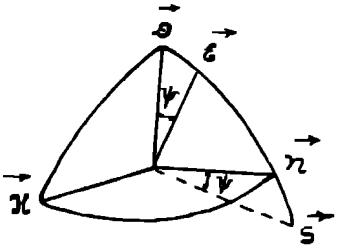


FIGURE 3

Expression for cos Ψ shows that all the conditions reduce to the single condition $\Psi = \frac{1}{2}\pi$, for then $\mu^2 = 0$ (Appleton-Hartree condition), E becomes parallel to N as suggested by Bose and S becomes perpendicular to N (Booker's condition)

Thus the different postulates about reflexion of waves are identical

NATURE OF THE PROBLEM AT PRESENT

The above discussions have unified the different works on the problem of the propagation of the electromagnetic wave in the ionosphere as treated from the standpoint of the ray method. The shortcomings of this method have already been pointed out

Recently, several workers have attacked the problem from the standpoint of wave treatment, though the inferences drawn therefrom are not substantially conclusive Hartree (1932) attempted to solve the wave equation by assuming a linear ion-barrier, which does not at all tally with the actual case Appleton, Naismith & Ingram (1939) have quoted a result of Booker's, who has given a solution of the wave equation for a parabolic layer, valid for frequencies near the critical value, but invalid for other waves Wilkes (1940), Rawyer (1939), Rydbeck (1940) and others have attempted to solve the wave equation with different hypothetical ion-barriers which do not fit the actual case Saha & Rai (1937) attempted to solve for a general barrier, using B K W approximate method and the shortcomings of this method are pointed out by the authors themselves The complete forms of the wave equation have been deduced by Saha & Banerjea (1945) by using the principal axes $HY'Z'(XYZ)$, and they have the form

$$\nabla^2 E_x + \frac{p^2}{c^2}\left(1 - \frac{p_0^2}{p^2 - \imath\nu p}\right) E_x = 0,$$

$$\nabla^2 (E_y \pm \imath E_z) + \frac{p^2}{c^2}\left(1 - \frac{p_0^2}{p^2 - \imath\nu p \mp p p_h}\right)(E_y \pm \imath E_z) = 0.$$

Here E_x, E_y, E_z have been written for E_H, $E_{y'}$, $E_{z'}$ for the sake of simplification of symbols.

From these equations, we see clearly that on entrance into the ionosphere, the three components E_x, $E_y + iE_z$, $E_y - iE_z$ travel with different velocities and absorption, depending on p_0^2, ν, p_h

The wave equations are of the type, commonly known as the Riccatti equation, which often reduces to integrable equations like Bessel, Legendre, hypergeometric, Airy's integral, Whittaker's integral for special functional expressions for the variable refractive index Some of these special forms have been used by the authors mentioned above, though the general cases have still not been attempted In fact, the general case of the Riccatti's equation cannot be solved analytically but it can often be reduced to approximate integrable forms, as in B K W approximate processes The analytical justification of such approximations are still lacking as is evident from the discussions of Langer (1937) The alternative method is to solve the problem by the method of integral equation, as the problem is much too similar to Liouville's problem in integral equations

To fit the actual ionospheric case, however, other less accurate but simpler methods are available For the E region, the ion-barrier is given as a function of height by Chapman (1931) and Pannekoek (1929), and the results though originated from quite different starting points are, as shown by Saha & Rai (1937), concordant. The nature of this function has also been tested experimentally to fit the E-region. So the Riccatti equation can be put in this form, and its solution can be obtained analytically for different ranges, though no functional solution holds good for the entire range Then the different solutions can be made to fit each other by the common method of matching In this way a less analytic but more useful solution can be obtained, which when properly adjusted to fit the boundary conditions will link up the strength of the received signal with the circuital constants and the wattage transmitted

The alternative numerical method is to solve the equation with Chapman barrier with the help of Bush's differential analyser and then to fit a group of analytical functions with the actual solutions. By this method also we can arrive at expressions like reflexion coefficient, etc., which are directly observable Attempts are made along both lines by the author

The work was undertaken at the suggestion of Professor M. N. Saha, F R S., to whom I express my sincere thanks for constant help and guidance I have been the recipient of a Research Scholarship from the Radio Research Committee of the Council of Scientific and Industrial Research, and wish to express my thanks to the Council for facility to work and permission to publish this paper.

References

Appleton 1932 *J Instn Elect Engrs*, **71**, 646
Appleton 1935 *Reports on progress in physics*, **2**.
Appleton & Barnett 1925 *Proc Roy Soc* A, **109**, 621
Appleton & Ratcliffe 1928 *Proc Roy Soc* A, **117**, 576
Appleton, Naismith & Ingram 1939 *Proc Phys. Soc* **51**, 243
Bajpai & Pant 1937 *Sci & Cult* **2**, 409.
Balfour Stewart 1878 *Encyc Brit* **16**, 181.
Banerjee 1945 *Indian J Phys* **19**, 75
Berkner & Wells 1937 *Terr Magn* **42**, 73
Booker 1936 *Proc Roy Soc* A, **155**, 235
Bose 1938 *Indian J. Phys.* **21**, 121
Breit & Tuve 1926 *Phys Rev* **28**, 554
Bruewitch 1934 *Nature*, **133**, 175
Chapman 1931 *Proc Phys Soc* **43**, 26.
Colwell, Friend & Hall 1936 *Rev Sci Instrum* **7**, 420
Darwin 1924 *Trans Camb. Phil Soc* **23**, 137.
Darwin 1934 *Proc Roy Soc.* A, **146**, 17.
Goubau & Zenneck 1930 *Phys Z* **31**, 333
Green 1932 *Bull Radio Res. Bd, Aust.*, no 59.
Harang 1937 *Terr Magn.* **41**, 143
Hartree 1931 *Proc Camb Phil Soc* **27**, 143
Langer 1937 *Phys Rev* **51**, 669
Pannekoek 1929 *Proc Acad Sci. Amst* **29**, 1165
Rai 1937 *Proc Nat Inst. Sci India*, **3**, 307.
Rawyer 1939 *Ann Phys., Lpz*, **35**, 385
Rydbeck 1940 *J Appl Phys* **13**, 577.
Saha & Banerjea 1945 *Indian J Phys* **5**, 159
Saha & Rai 1937 *Proc Nat. Inst Sci India*, **4**, 319.
Saha, Rai & Mathur 1937 *Proc Nat Inst Sci. India*, **4**, 53
Toshniwal 1935 *Nature*, **135**, 1071
White 1931 *Proc Camb Phil Soc* **27**, 445
Wilkes 1940 *Proc. Roy Soc* A, **175**, 143.

The intermetallic compound phases of the system aluminium-manganese-zinc

By G V Raynor and D W. Wakeman

Department of Metallurgy, The University of Birmingham

(*Communicated by C Sykes, F.R.S.—Received* 17 *September* 1946)

[Plates 3 and 4]

The system aluminium-manganese-zinc has been examined in the range 0 to 95 % of zinc, and 0 to 3 % of manganese Attention was directed only to the constitution of the alloys above the solidus Using micrographic and X-ray methods, and the chemical analysis of crystals separated from slowly cooled alloys, it has been shown that, according to composition, the phases $MnAl_6$, T_1, $MnAl_4$, T'_2, T_3 and $MnAl_3$ may crystallize as primary constituents Both $MnAl_6$ and $MnAl_4$ dissolve small quantities of zinc, the phases T_1, T_2 and T_3 are ternary compounds The phase T'_1 is characterized by a ratio of four aluminium atoms to one of combined solutes, and an electron atom ratio of 1 85, calculated on the basis of the Pauling theory of transitional metals According to this theory, transitional metal atoms have vacancies for electrons in their atomic orbitals, and the present experiments in conjunction with earlier work suggest that these may be filled up as a consequence of alloy formation The phases T'_2 and T_3 may be represented respectively by the formulae Mn_2ZnAl_2 and $(Mn\ Zn)_5Al_{11}$. $MnAl_3$, which can dissolve small quantities of zinc, enters into equilibrium at a ternary eutectic (Zn 95 %, Mn 0 05 %, 378° C) with the primary solid solutions in zinc and aluminium respectively

1. Introduction

In the course of work on the equilibrium relations in the systems aluminium-manganese-nickel and aluminium-manganese-copper (Raynor 1944*a*, *b*), it was shown that the ternary compounds which were formed, and which entered into equilibrium with the primary aluminium-rich solid solution, could be regarded as based upon a common plan In the latter system, the ternary compound extends from copper 7 8, manganese 12 1, aluminium 80 1 atomic % to copper 6 9, manganese 15 0, aluminium 78·1 atomic %, and appears to be based upon the arrangement $(Cu\ Mn)Al_4$ The composition which is in equilibrium with both $CuAl_2$ and the primary aluminium-rich solid solution corresponds almost exactly with the arrangement $Cu_2Mn_3Al_{20}$ In the aluminium-manganese-nickel system the corresponding ternary compound is also based upon the arrangement $(Ni\ Mn)Al_4$, and experiment suggests that it is almost of the fixed composition $Ni_4Mn_{11}Al_{60}$.

The variations in the compositions characterizing these two ternary compounds have been interpreted (Raynor 1944*b*) in terms of the Pauling theory of the transitional metals (Pauling 1938) According to this theory, 5·78 electrons of the 3*d* and 4*s* quantum groups of the atoms of the transitional metals of the first long period may be regarded as responsible for cohesion, while the remainder occupy 'atomic orbitals', of which 2·44 are available For the transitional metals of the first long

period, therefore, vacancies for electrons exist in the atomic orbitals If it is assumed that these vacancies may be filled by electrons absorbed from the structure as a whole when the transitional metal forms a minor constituent of a binary or ternary alloy, the electron atom ratios for the ternary compounds in the aluminium-manganese-copper and aluminium-manganese-nickel systems respectively may be evaluated as 1·86 and 1·84 * Further evidence in favour of this absorption of electrons by transitional metals has been obtained by a consideration of the lattice spacings of the primary solid solutions of chromium and manganese in aluminium (Raynor 1945)

It appears possible, therefore, that in this type of aluminium-rich ternary alloy the ternary compound which exists in equilibrium with the primary solid solution tends to be based upon an electron atom ratio of approximately 1 85, and at the same time to satisfy the condition that there should be four atoms of aluminium to one atom of combined solutes The latter condition must clearly depend upon size relationships between the atoms composing the structure.

In order to gain further information with regard to this possibility, the system aluminium-manganese-zinc has been investigated In this case the substitution of zinc for nickel or copper leaves the atomic size relationships relatively unaffected; the valency factor, however, increases by unity on passing from copper to zinc The system would be expected, therefore, to contain a ternary compound based upon the ratio of four aluminium atoms to one of combined solutes, but with the composition so adjusted as to preserve the electron atom ratio at approximately 1 85. Assuming that aluminium and zinc respectively contribute to the structure, 3 and 2 electrons per atom, and that manganese, in accordance with Pauling's theory, has vacancies for 3 66 electrons per atom, the predicted composition of the expected ternary compound is.

Manganese 28 01, Zinc 6·40, Aluminium 65 59 weight %

2 Previous work

(*a*) *The binary systems*

The equilibrium diagrams of the binary systems aluminium-zinc and aluminium-manganese are satisfactorily established. Figure 1 shows the diagram for the aluminium-zinc alloys (Raynor 1943) according to the results of the most reliable work The liquidus curves fall smoothly from the freezing points of the pure metals to a eutectic at 95 % of zinc and a temperature of 382° C The solid solubility of aluminium in zinc is limited, but that of zinc in aluminium is extensive Between the composition limits 31 6 and 78 % of zinc, and the temperature limits 365° and 275° C, this solution splits up into a mixture of two face-centred cubic phases.

Figure 2 shows the equilibrium diagram of the aluminium-manganese system, based upon the results of Dix, Fink & Willey (1933), Phillips (1943), Butchers &

* The aluminium-manganese-copper ternary phase has a range of homogeneity The electron-atom ratio is calculated for the composition which is in equilibrium with $MnAl_6$ and the aluminium rich solid solution

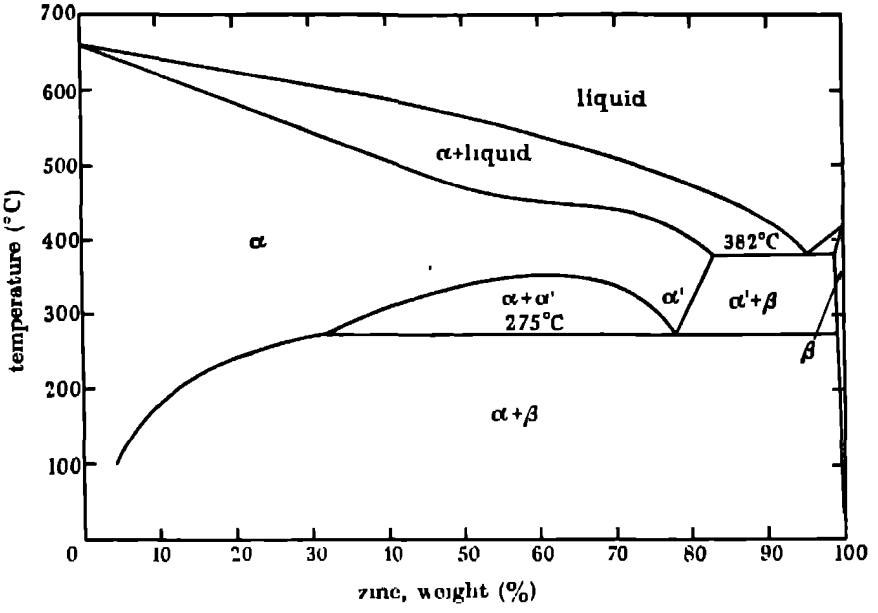


FIGURE 1

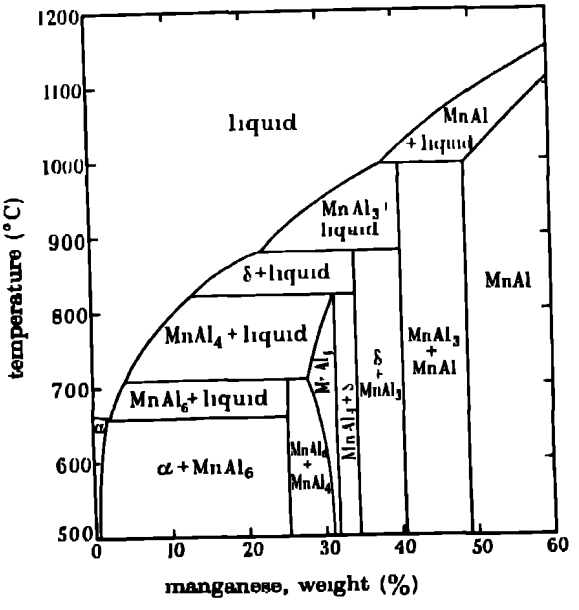


FIGURE 2

Hume-Rothery (1945) and Koster & Bechthold (1938) A eutectic point, at which liquid$\rightleftharpoons\alpha+MnAl_6$, occurs at 1 95 % of manganese and a temperature of 658 5° C, the solid solubility of manganese in aluminium is limited, particularly at low temperatures $MnAl_6$ is formed by a peritectic reaction at 710° C from $MnAl_4$, which in turn is formed peritectically from a δ-phase containing approximately 34 % of manganese Further peritectic reactions involve the formation of δ from $MnAl_3$, and of $MnAl_3$ from MnAl, which is a phase with a wide range of homogeneity extending from 49 to 67 % of manganese Each of the four peritectic reactions tends to remain incomplete on cooling, so that structures which are not in equilibrium are readily obtained.

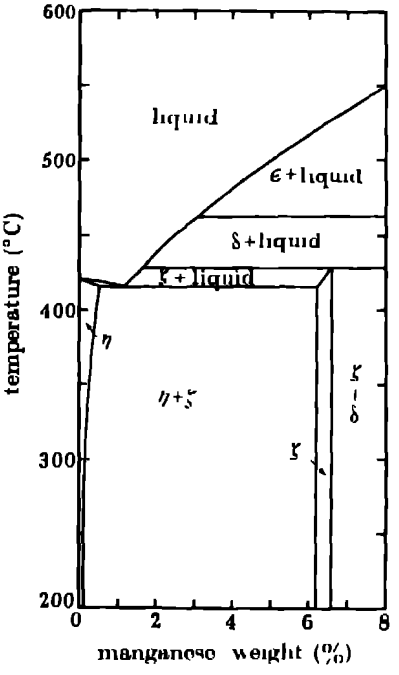


FIGURE 3

The zinc-rich portion of the zinc-manganese system as given by Schramm (1940) is shown in figure 3 A eutectic is formed at 416° C between the zinc-rich solid solution, which can hold 0·5 % of manganese in solid solution at this temperature, and a monoclinic ζ-phase containing 6·2 % of manganese Further peritectic reactions, involving the ϵ- and δ-phases of this system, occur at 462° and 428° C, but these are of little importance for the present work

(*b*) *The ternary system*

No previous work on the aluminium-rich alloys of the system aluminium-manganese-zinc appears to have been published. Experiments on the equilibrium relations in the zinc-rich corner of the phase diagram have, however, been reported by

Gebhardt (1942), and are summarized in figures 4 and 5, which show respectively the projection of the surfaces of primary separation, and the equilibrium relations in the solid state at 350° C At U_T (zinc 99·15, aluminium 0 55, manganese 0 3 %; 412·5° C) the liquid and the binary zinc-manganese ζ-phase react together to form the zinc-rich solid solution and a phase denoted V A ternary eutectic involving the aluminium-rich and zinc-rich solid solutions and the compound V is formed at E_T (378° C). The composition of the eutectic point is given as 95 % of zinc and 0 04 % of manganese This point is therefore almost coincident with the binary aluminium-zinc eutectic. Figure 5 shows that at temperatures close to that of the eutectic the zinc-rich solid solution can take up approximately 0 5 % of manganese and 1·0 % of aluminium, while the face-centred cubic solid solution based on aluminium can take up approximately 83·0 % of zinc. The composition of this latter phase which enters into equilibrium with the zinc-rich solid solution and the V-phase, however, corresponds with only 0·05 % of manganese The composition of the V-phase was not accurately established, it had the characteristics of an intermetallic compound, and extrapolation of the sides of the three-phase triangle inside which it is in equilibrium with the aluminium-rich and zinc-rich solid solutions suggested that it contained little zinc, and approximately 30 % of manganese

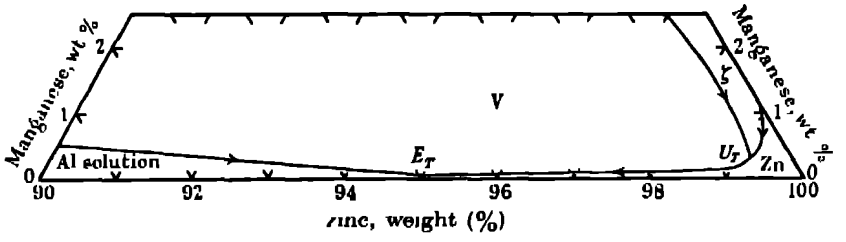


FIGURE 4

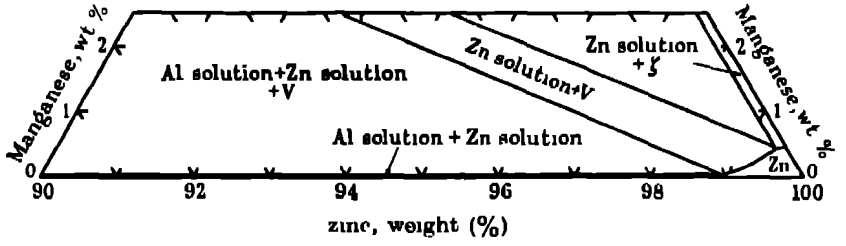


FIGURE 5

From a comparison of figures 1, 2 and 4, it is clear that the ternary aluminium-manganese-zinc equilibrium diagram must contain a very long eutectic valley, stretching from the binary aluminium-manganese eutectic to the ternary eutectic point at 95 % of zinc In this respect the system differs from those previously examined, this is of importance with respect to the number of phases which may enter into equilibrium with the aluminium-rich solid solution.

3. Experimental methods

Preliminary experiments, involving the annealing of chill-cast alloys having compositions within the range 0 to 5 % of manganese and 0 to 60 % of zinc, showed that equilibrium was very difficult to attain within a reasonable time of annealing. Many alloys, after annealing periods of more than 3 weeks, showed four phases, in the form of hard particles of three different types, in considerable relief, set in a matrix of aluminium-rich solid solution Experiments indicated that these phases were all present in the chill castings, and that the elimination of the unstable phase or phases by subsequent annealing was very slow In order, therefore, to obtain information with regard to the phases which might be present, it was decided to investigate the system by the method of slow cooling from the melt, with subsequent micrographical examination Difficulties encountered made it essential to extract and examine the primary crystals deposited in the alloys

(*a*) *Materials used*

The materials used in the preparation of the alloys were as follows

(i) Pure zinc The only elements detected spectrographically, and their approximate amounts, were lead 0 0001 %, cadmium 0·00005 % and calcium 0 0001 % For certain preliminary alloys, zinc of 'Crown Special' grade was used It contained lead 0 0007 %, cadmium 0 0010 %, iron 0·0010 % and copper 0 0002 %. These materials were presented by the Imperial Smelting Corporation, Ltd., of Avonmouth, to whom the authors express their sincere gratitude.

(ii) Super-purity aluminium, kindly presented by the British Aluminium Company, Ltd

(iii) Aluminium-manganese master alloys, of which one, containing 9 52 % of manganese, was presented by the Mond Nickel Company, and the other, containing 5 25 % of manganese, was presented by the British Aluminium Company, Ltd. These master alloys were prepared from pure materials, and the authors' thanks are due to the laboratories in which they were produced

(*b*) *Preparation of alloys*

Alloys containing less than 60 % of zinc were prepared in quantities of 50 to 100 g Those containing 60 to 95 % of zinc were made up in quantities of 100 g In all cases weighed-out quantities of aluminium and the aluminium-manganese master alloy were melted together in a crucible lined with an alumina-fluorspar mixture, using a small electric crucible furnace. After thorough stirring with a sintered alumina stirring rod, the requisite quantity of pure zinc was added and quickly stirred into the melt, which was then allowed to cool slowly in the furnace The rates of cooling varied somewhat at different stages of the research, but were in general of the order of 1 to 2° C/min

When cold, the cylindrical ingots were sectioned longitudinally, and one section ground, polished and examined metallographically. The results of the micrographic

examination were difficult to interpret, no etching reagent was found which would satisfactorily differentiate between different intermetallic compound crystals, and the constitution could only be judged qualitatively from the general appearance of the microsections.

(c) *Extraction of primary crystals*

Primary, and secondary, crystals present in the alloys were extracted by an electrolytic method Each ingot was made the anode of an electrolytic cell (figure 6), of which the cathode was a nickel strip and the electrolyte a mixture of dilute hydrochloric and nitric acids Under these conditions, using a current density of approximately 0·1 amp./sq in , and a potential difference of 8 V, the solid solution matrix was slowly dissolved away The intermetallic compound crystals present fell away to the bottom of the cell When sufficient residue for examination had collected, the electrolysis was stopped, and the residue well washed with 50 % nitric acid, followed by several washings with distilled water After rinsing with alcohol, the residue was dried by gentle heating.

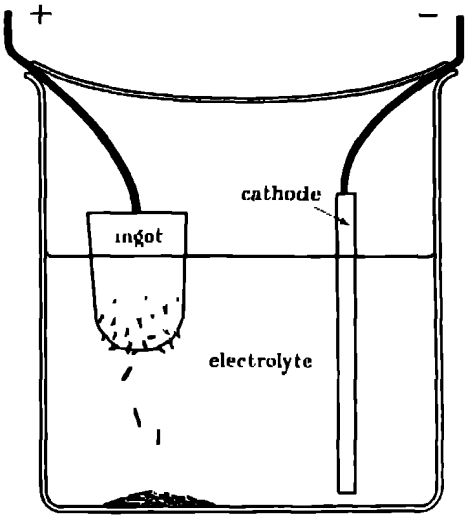


FIGURE 6

All residues were first examined under a binocular microscope Details of these experiments are given below. In all cases the extracted crystals were clean and apparently unattacked by the reagents used for extraction Crystal faces were often well developed and of very high reflectivity

In no case, under any conditions of cooling which were attempted, were homogeneous residues obtained Examination under the microscope showed that the extracted residues were mixtures of two, or sometimes three, different intermetallic compounds, with clearly differentiated habits. In order to identify these, crystals of each crystallographic type were hand-picked, using forceps, from the mixture,

and submitted to chemical analysis. In critical cases X-ray examination was also used Analysis confirmed the differentiation of the various phases on the basis of their observed habits.

(*d*) *Analytical method*

Since the hand sorting of extracted residues was somewhat laborious, analytical methods capable of giving accurate results with small quantities of material were explored In general, 0 2 to 0·3 g. of residue could be collected from the mixtures by picking out good crystals about whose crystallographic habit there could be no reasonable doubt

The analytical method finally adopted, after repeated tests on synthetic solutions containing aluminium, manganese, and zinc in approximately the same concentration as that to be expected during the assay of extracted crystals, was as follows

The crystals were dissolved in the minimum quantity of 50 % sulphuric acid, and the clear solution made alkaline with caustic soda until the hydroxides of zinc and aluminium had redissolved Hydrogen peroxide was added, and the mixture boiled to decompose excess of this reagent The supernatant liquid was filtered from the precipitate of hydrated manganese dioxide using a sintered glass crucible To determine the manganese, the precipitated dioxide was dissolved in nitric acid containing a little sulphurous acid, and the solution diluted and boiled The solution was oxidized with sodium bismuthate, filtered, and the manganese in the filtrate estimated in the usual way by the addition of an excess of ferrous ammonium sulphate and back-titration with potassium permanganate solution The zinc was estimated in the acidified filtrate from the manganese precipitate by direct titration with standard potassium ferrocyanide solution (containing a little potassium ferricyanide), using diphenylamine as an internal indicator All solutions employed were regularly standardized, and reproducible results were obtained throughout the research

(*e*) *X-ray examination*

Certain specimens of extracted crystals were submitted to X-ray examination For this purpose they were crushed to powder in an agate mortar, and mounted

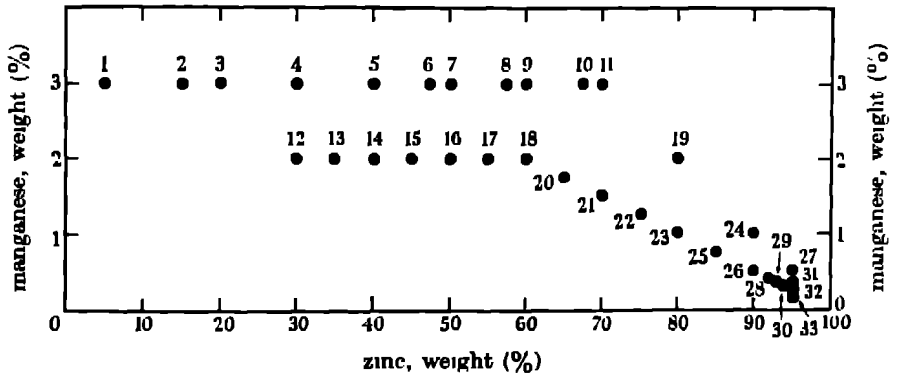


FIGURE 7

on hairs in the usual manner with canada balsam dissolved in xylol. Specimens were exposed, in a cylindrical Debye-Scherrer camera, to iron $K\alpha$ radiation from a Metropolitan-Vickers demountable X-ray machine.

In the course of the work, thirty-three alloys were examined. Their compositions are shown diagrammatically in figure 7. In order to facilitate description of individual alloys, compositions will be referred to in terms of the weight percentages of manganese and zinc in the alloy. Thus, the term 'alloy 2·0/80' refers to an alloy containing 2 0 % of manganese and 80 % of zinc

4 The microstructures of the alloys

As noted above, all alloys were carefully examined micrographically before residue extraction In general, these examinations were not informative, owing to the fact that it was extremely difficult to distinguish between the intermetallic compound phases with any certainty in either the unetched or etched conditions The hard crystals, being set in a much softer matrix, were always seen in considerable relief, which was difficult to avoid In addition, the large difference in electrochemical characteristics of the particles and the matrix led to all etching reagents giving vigorous attack on the matrix, leaving the hard particles relatively unaffected. No marked differences in relief between various types of hard particles were observed Distinctions in the shapes of particles in the microsections were noted, but subsequent examination of the extracted residues showed that such distinctions could be quite misleading No detailed description of the microstructures observed will therefore be given, except in so far as they assist the interpretation of the results and the construction of the diagram showing the surfaces of primary separation

The first point to be noted is that the crystals in the microsections of alloys 3 0/20, 3·0/30 and 3·0/40 were mostly homogeneous but occasionally included small spines of a slightly darker material (figure 8, plate 3) Since these crystals were subsequently shown to be $MnAl_{6}$, it was concluded that these compositions were just on the edge of the field in which the phase $MnAl_4$ deposits as primary crystals, but that these small first-deposited crystals were quickly surrounded, by peritectic reaction, by thick sheaths of $MnAl_6$, which then continued to separate as the primary constituent until the temperature was reached at which $MnAl_6$ and the primary solid solution separate together

No further evidence of peritectic sheaths round cores of a different material was obtained in alloys containing less than 00 % of zinc, with the exception of alloys 3·0/70 and 2 0/80 Alloy 3 0/70 showed primary needles, some of which had spines of darker material, and secondary plates, some of which, in cross-section, had the appearance of needles. A portion of the microstructure is shown in figure 9, plate 3. Alloy 2·0/80 showed primary crystals of a dendritic type (figure 10, plate 3), together with plate-like material A few of the dendritic crystals, however, possessed small spines of darker material. In both these cases, therefore, the alloys lie on the edge of a region in which some unidentified phase deposits, which is subsequently trans-

formed by peritectic reaction into the phase which predominates in the actual microstructure.

The alloys containing 90 and 95 % of zinc will be discussed below with regard to the nature of the ternary eutectic

Figures 8 to 13, plate 3 show a representative series of microstructures in which the phases have been identified as a result of the experiments described in the following sections

5 General description of extracted residues

It was not possible, from the micrographic examination, to decide how many phases existed in equilibrium with the primary solid solution, or to assess the limits of their occurrence In the examination of extracted residues by means of the binocular microscope, however, seven distinct forms of intermetallic compound crystal were observed

The extracts from alloys 3 0/5 and 3 0/15 consisted entirely of long needle-shaped crystals in the form of rectangular prisms (figure 14, plate 4) These crystals were exactly of the form associated in earlier work (Raynor 1944*b*, Raynor & Hume-Rothery 1943) with the compound $MnAl_6$

The extract from alloy 3 0/20 was also predominantly of this type, but contained in addition a small amount of a different type The latter crystals were small, regular in shape, and approximately equi-axed as shown in figure 15, plate 4 Extracts from alloys 3·0/30 and 2 0/30 showed an increasing amount of this constituent in addition to the suspected $MnAl_6$ crystals, while those from alloys 2·0/35 and 3 0/40 consisted almost entirely of the regular-shaped crystals, with a small amount of the suspected $MnAl_6$ phase. In addition, the extracts from 2 0/35 and 3·0/40 contained a few very fine long needles, which were clearly of a different type from those which were suspected to be $MnAl_6$. The thin needles are shown in figure 16, plate 4.

This behaviour was interpreted as indicating that in these alloys $MnAl_6$ was the primary constituent, but that a secondary separation of the equi-axed crystals took place. This view was confirmed by the fact that, in the extracts from all alloys between this range of compositions and 65 % of zinc, no suspected $MnAl_6$ crystals were observed The residues consisted largely of the equi-axed crystals, which were somewhat increased in size, together with the fine needles, which increased in amount as the zinc content of the parent alloy increased In this range of compositions, therefore, the equi-axed crystals form the primary constituent, and are followed by a separation of fine needles.

At zinc contents greater than 65 %, none of the regular equi-axed crystals were observed The extract from 3·0/67·4 consisted almost entirely of the thin needles, while those from alloys 3 0/70 and 1 5/70 contained large amounts of these needles. It was concluded that in this range the thin needles were the primary phase. It was observed that the needles were approximately hexagonal in cross-section, and

microstructures of these alloys before residue extraction showed areas which were very similar in appearance to those containing hexagonal $MnAl_4$ crystals published by Phillips (1943).

In addition to primary needles, the extracts from alloys 3 0/70 and 1·5/70 contained plates, which were very thin, of high reflectivity, and sometimes considerable area These plates tended to be rectangular in shape, as illustrated in figure 17, plate 4. The amount of the thin plate-like crystals increased considerably in the extract from alloy 1·25/75, but the primary crystals were still of the needle type. This evidence suggests that primary crystallization of the needles is followed by a secondary deposition of the thin plates With the alloys employed, the region in which the plates separate as primary crystals was not observed

The extracts from alloys 2 0/80, 1·0/80, 0 75/85 and 0 5/90 all showed primary crystals of a completely different type from any observed in alloys less rich in zinc. The crystals were of a peculiar dendritic form, as indicated in figure 18, plate 4, with the surfaces parallel to the directions of the dendritic arms almost parallel to each other Secondary crystals were also present in these extracts and took the form of thin shiny plates of a different type from those previously observed These thin plates (figure 19, plate 4) had a high reflectivity and were seen under the microscope to possess characteristic hexagonal markings on their surfaces. They were very fragile and easily broke into pieces during the sorting operations In the unsorted residue, however, many complete plates were seen to be approximately hexagonal in shape

The residue from alloy 1·0/90 was, however, of a different type, this contained no dendritic crystals, but was composed of a mixture of two types of flat plates, of which one was easily recognizable as the same material as the shiny plates met with in the 0·5/90 alloy The other crystals appeared to be of a new type, and the constitution of this alloy is discussed more fully below in connexion with the nature of the ternary eutectic

In this work, therefore, the existence of several types of primary crystals, according to the composition of the ternary alloy, was recognized. The persistence with which two or more types of crystal are deposited from the melt is explained by the very long eutectic valley extending from the aluminium-manganese axis almost to the aluminium-zinc axis at 95 % of zinc Reference to figure 20, which shows schematically the forms of primary surfaces in a hypothetical equilibrium diagram containing a long eutectic valley, indicates that, whatever primary phase separates, deposition of one or more other phases will follow An alloy of composition X, on cooling to the liquidus surface, will deposit the primary A phase, and the composition of the remaining liquid will, on further cooling, move towards the eutectic valley pq. When the composition of the liquid reaches this line, simultaneous deposition of α and A occurs, and the composition of the liquid moves along the line pq towards q When the point r has been passed, α and B separate together; after passing the point s, α and C separate, and so on. The process continues as long as there is any liquid left; if cooling were sufficiently rapid for solidification to end

at the ternary eutectic q, the alloy X would contain, in addition to primary A, crystals of the B, C and D phases Under the conditions used in the experiments, the rate of cooling was sufficiently slow to limit the number of intermetallic compound phases present to two or sometimes three.

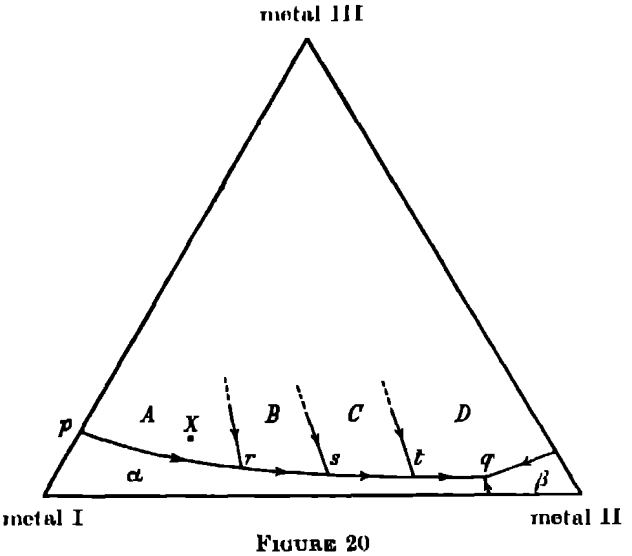


FIGURE 20

6. ANALYTICAL RESULTS

The results of the chemical analysis of thirty-five pure specimens, hand-sorted as previously described, are given in table 1, which is self-explanatory, and are plotted in figure 21 In some cases, where the composition of a phase had been

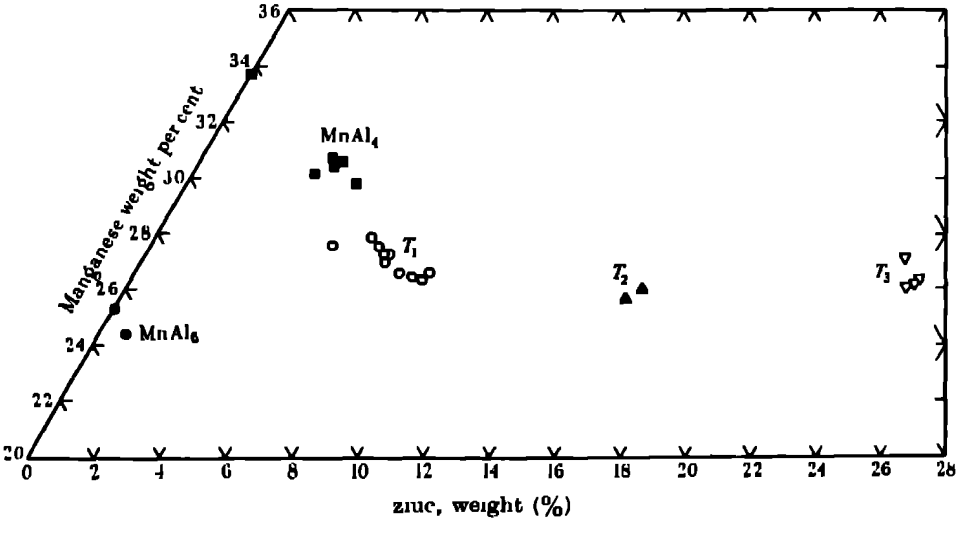


FIGURE 21

established by the determination of both manganese and zinc, the manganese alone was estimated in order to confirm the identification of the crystals in other extracts

Reference to figure 21 shows that the compositions of the crystals extracted from alloys containing less than 95 % of zinc fall into five distinct groups. The analyses

TABLE 1

alloy	type of extracted crystal	Mn %	zinc %
3 0/5	prismatic needles	—	0 36
3 0/30	prismatic needles	23 88	—
2 0/30	prismatic needles	24 41	0 78
2 0/35	regular equi-axed	27 90	6 51
3 0/40	prismatic needles	23 95	—
	regular equi-axed	27 3	—
		27 6	5 5
2 0/40	regular equi-axed	27 55	6 92
	thin needles	30 2	—
2·0/45	regular equi-axed	27 29	7 4
	thin needles	30 2	3 66
3 0/47 4	regular equi-axed	27 97	6 48
3 0/50	regular equi-axed	27 18	—
		27 3	7 18
	thin needles	30 6	4 3
2 0/50	regular equi-axed	26 98	7 4
	thin needles	30 1	—
2 0/55	regular equi-axed	26 65	8 01
3 0/60	regular equi-axed	26·37	8 8
	thin needles	30·2	—
2 0/60	regular equi-axed	26 49	8 45
	thin needles	30 68	4 01
1·75/65	regular equi-axed	26·65	8·92
	thin needles	30 53	4 07
3 0/67·4	thin needles	30 3	—
3 0/70	thin needles	30 75	—
1·5/70	thin plates	25 70	15·37
1 25/75	thin needles	29 75	5·15
	thin plates	26 06	15·70
2 0/80	dendritic crystals	26 01	23 81
1 0/80	dendritic crystals	27 14	23 15
0·75/85	dendritic crystals	26 35	24·03
0 5/90	dendritic crystals	26 13	23 92
0 25/95	plates with hexagonal markings	38 88	0 5*
0 35/95	rectangular plates	50 23	none detected*

* See § 8.

confirm that the prismatic needles are $MnAl_6$, which can take up small amounts of zinc into solid solution The regular equi-axed crystals, the thin plates, and the dendritic crystals clearly correspond with three different ternary compounds, which have been denoted T_1, T_2 and T_3 respectively The analysed compositions of the thin needles suggest that they correspond with $MnAl_6$, which has zinc in solid solution. Further reference is made to this point in §§ 7 and 10.

7. X-RAY RESULTS

In order to confirm the indications of the analytical and micrographic experiments, a series of X-ray diffraction patterns was obtained Comparison of the pattern given by the rectangular prism crystals from alloy 2 0/30 with that obtained from a sample of pure $MnAl_6$ extracted from a binary aluminium-manganese alloy showed without doubt that these crystals were correctly identified as $MnAl_6$

Diffraction patterns given by pure $MnAl_4$ crystals were similarly compared with that given by a typical specimen of the thin needles, separated from alloy 2 0/35 The patterns obtained were very closely similar in the general arrangement of the lines The diffraction lines in the films obtained from the needle-shaped crystals showed some very slight differences in position when compared with the corresponding $MnAl_4$ lines, this may easily be accounted for by the presence of zinc in the extracted crystals The general agreement of the patterns suggests clearly that the needle-shaped crystals correspond with a solid solution of zinc in the binary aluminium-manganese compound $MnAl_4$

Photographs were also taken of T_1 crystals from alloy 3 0/47·4, T_2 crystals from alloy 1 25/75, and T_3 crystals from alloy 2·0/80 These showed patterns which, as expected, were characteristically different from each other, and from those of the phases of the binary aluminium-manganese system These photographs, owing to relatively high general background scattering, were not suitable for reproduction It is hoped, however, to present a more detailed X-ray study of these phases later

8 EQUILIBRIUM IN THE REGION OF THE EUTECTIC

As explained in § 2 the ternary eutectic, according to the work of Gebhardt, involves equilibrium between the aluminium solid solution, the zinc solid solution, and a compound V containing little zinc and approximately 30 % of manganese. In an effort to characterize the phase V more closely, slowly cooled alloys with compositions in the neighbourhood of the eutectic were examined. As already indicated, alloy 1 0/90 gave a residue composed of two types of flat plates, one roughly rectangular in shape, and the other of irregular shape and high reflectivity, similar to those observed in other specimens. The microstructure of this alloy showed that many of the crystals were cored, thus rendering the extract unsuitable for accurate analysis Alloy 0 5/90 was free from these defects, and the extract consisted of the dendritic T_3 phase (see table 1) together with the irregular shiny plates These results suggest that, although the primary crystals in alloys 1 0/90 and 0 5/90 were different, both were followed by the crystallization of the thin shiny plates, in the former case by a peritectic reaction, and in the latter case by secondary deposition as explained in § 5 Alloy 0 5/95 showed a microstructure similar to that of alloy 1·0/90, and a similar extract was obtained, which was, because of the coring observed, unsuitable for analysis An alloy of composition 0 25/95, however, showed a microstructure in which the crystals were almost entirely free from any peritectic inclu-

sions, the extracted residue contained a few of the rectangular type of plate, and a quantity of the thin shiny material An alloy containing slightly more manganese (0·35 %) contained the same two types of crystal, but with the amounts reversed It appeared clear, therefore, that in both cases the rectangular type of crystal separated first, followed by the shiny plates The ternary eutectic therefore involves equilibrium between the latter and the two solid solutions Confirmation of this view was given by the extract from alloy 0·15/95, which, though very small in amount, consisted entirely of the thin shiny plates

As shown in table 1, analysis of the thin irregular plates proved them to contain 38·88 % of manganese, and 0·5 % of zinc, while the rectangular plates contained 50·23 % of manganese Reference to figure 2 makes it very probable that the former are to be identified with the phase $MnAl_3$, which occurs at 40 % of manganese in the binary diagram, and the latter with the phase MnAl, the aluminium-rich boundary of which occurs at approximately 50 % of manganese.

It is therefore confirmed that the intermetallic compound phase which enters into the ternary eutectic contains very little zinc, and the experiments suggest that it is the phase $MnAl_3$ Equilibrium relations in this region are, however, complex, and with a small increase in manganese content above that of the ternary eutectic, MnAl may separate as primary crystals. The diagram given by Gebhardt (figure 4) has therefore to be modified, and his V-phase identified with $MnAl_3$. The form of the surfaces of primary separation above 90 % of zinc, as suggested by the present work combined with that of Gebhardt, is given in figure 22.

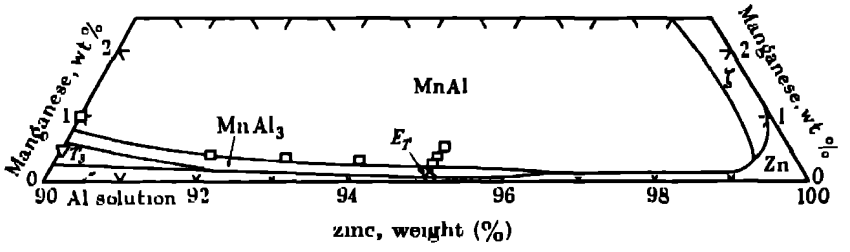


FIGURE 22

9. THE SURFACES OF PRIMARY SEPARATION

From the results described in the preceding sections, a diagram showing the form of the surfaces of primary separation may be constructed This is shown in figure 23, which is self-explanatory The boundaries of these regions of primary separation are only to be considered as approximate at the higher manganese contents, but the intersections of the boundaries with the eutectic valley are accurately established by the points plotted on the diagram.

Several features of interest are presented in figure 23. In the first place, the extension of the primary $MnAl_3$ and primary MnAl fields from the aluminium-manganese axis almost to the manganese-zinc axis, as suggested by these experi-

ments, is remarkable. It results in the crystallization, in the neighbourhood of 95 % of zinc and at temperatures of the order of 380° C, of compounds which separate only at temperatures above 900° C in the binary system aluminium-manganese

Again, the extension of the primary $MnAl_6$ field to the eutectic valley is of interest, and is easily understood as a consequence of the existence of this long valley, running nearly parallel to the aluminium-zinc axis of the ternary diagram

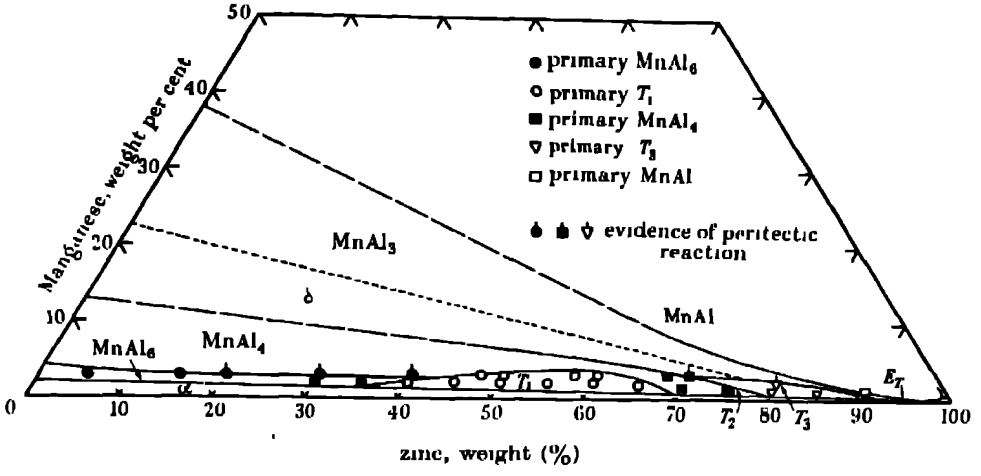


FIGURE 23

The present experiments provide no evidence that the field corresponding with the primary crystallization of the δ-phase of the binary aluminium-manganese system extends to the eutectic valley. The succession of phases deposited on passing along the eutectic valley towards the ternary eutectic is

$$MnAl_6 \rightarrow T_1 \rightarrow MnAl_4 \rightarrow T_2 \rightarrow T_3 \rightarrow MnAl_3.$$

The boundary between the primary $MnAl_3$ and primary δ field can therefore not extend to the eutectic valley, and has been tentatively inserted as a dotted line in figure 23

According to the experiments described, alloy 1·25/75 deposits primary $MnAl_4$, while alloy 1·0/80 deposits primary T_2 crystals. Between these two compositions, however, separation of the T_2 phase occurs, so that the region in which this phase crystallizes as the primary constituent must be confined to a narrow range of compositions between 75 and 80 % zinc, as shown.

10. Discussion

The results described in the preceding sections indicate that the equilibrium relations above the solidus in the system aluminium-manganese-zinc, with small amounts of manganese, are complex in nature No fewer than six different regions

of primary separation extend to the eutectic valley which joins the binary aluminium-manganese eutectic to the zinc-rich ternary eutectic. All six of the phases which crystallize in these primary fields may therefore be expected to exist, in the solid alloys, in equilibrium with the primary solid solution based on aluminium This somewhat unusual state of affairs is due essentially to the wide solid solubility of zinc in aluminium

(*a*) *The compound* $MnAl_6$

Up to a zinc content of approximately 35 % at 2 % of manganese, or 43 % at 3 % of manganese, the primary phase is $MnAl_6$, which can take up small quantities of zinc The composition of the $MnAl_6$ crystals which separated from alloy 2 0/30 shows that at least 0 78 % of zinc may be held in solid solution Since $MnAl_6$ will separate from alloys richer in zinc, it is improbable that this amount of zinc corresponds with the maximum solid solubility of zinc in $MnAl_6$ The maximum solubility is not likely, however, greatly to exceed this value The analysed composition of the crystals from alloy 2 0/30 corresponds with the formula $Zn_{0\cdot03}Mn_{0\cdot96}Al_6$, thus providing some evidence that zinc replaces manganese in the compound atom for atom.

(*b*) *The ternary compound* T_1

In a region from 35 to approximately 69 % of zinc, for manganese contents between 1 and 3 %, the ternary compound T_1 crystallizes as the primary constituent The composition of this compound is of particular interest, from the point of view of the theory referred to in § 1. As stated above the compound may be expected to possess a composition based upon a ratio of four aluminium atoms to one of combined solutes, and an electron atom ratio of 1·85

For calculation of electron atom ratios for structures containing transitional metals, it is assumed that each transitional metal atom absorbs from the structure as a whole sufficient electrons to fill up its 'atomic orbitals' to the maximum permissible extent According to the theory developed by Pauling (1938), the transitional metals of the first long period have the following vacancies per atom in their atomic orbitals.

Cr	Mn	Fe	Co	Ni
4 66	3 66	2 66	1 71	0 61

In the system aluminium-manganese-zinc, therefore, each aluminium atom contributes three electrons to the structure and each zinc atom contributes two; each manganese atom, however, absorbs 3 66 electrons from the structure as a whole

In figure 24, the line *ab* represents a constant ratio of four aluminium atoms to one of combined solutes, while the line *cd* represents a constant ratio of 1 85 valency electrons per atom of the structure, calculated as indicated above When the compositions of the compound T_1 as determined by analysis are also plotted, it is seen that, except for the residue from alloy 3·0/40 which may have been contaminated with a small amount of $MnAl_6$, the points representing extracts from alloys containing less than 55 to 60 % of zinc lie accurately on the of line constant atomic

ratio. The point representing the sample with the least zinc in it corresponds almost exactly with the intersection of the constant atomic ratio line and the constant electron : atom ratio line, and is therefore in excellent agreement with the predicted composition (§ 1) It is clear, therefore, that the formation of this compound depends upon the simultaneous satisfaction of two factors—an electron atom ratio of 1 85 and a ratio of four aluminium atoms to one atom of combined solute The compound is analogous to the ternary compounds observed in the aluminium-manganese-nickel and aluminium-manganese-copper systems.

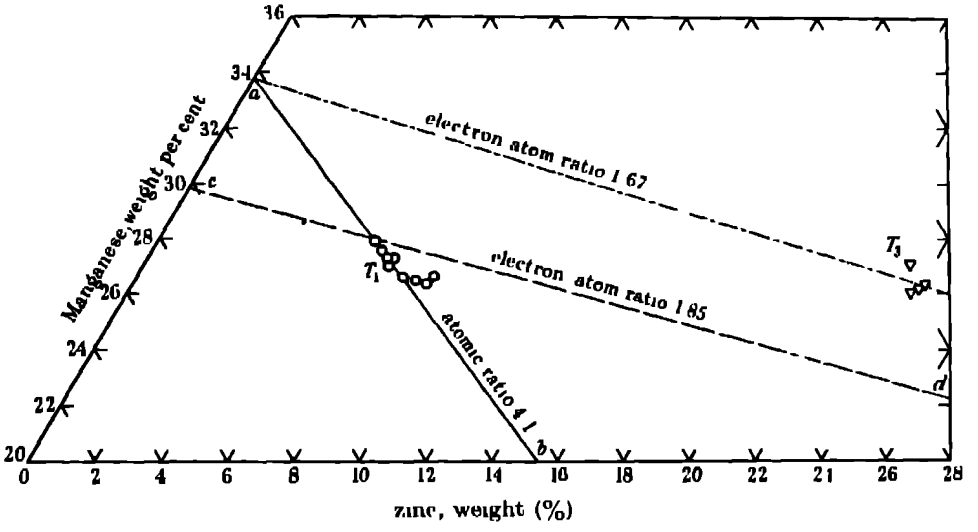


FIGURE 24

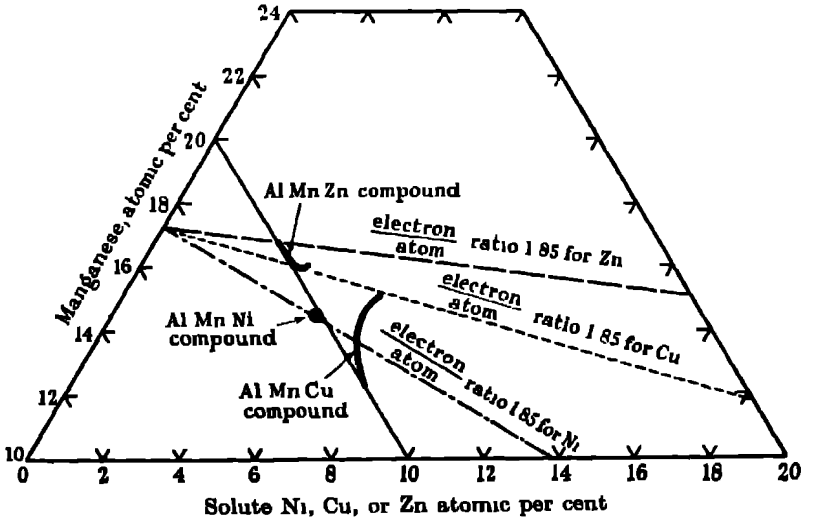


FIGURE 25

To facilitate comparison, in figure 25 the compositions of the corresponding ternary compounds in the alloys of aluminium and manganese with copper, nickel, and zinc are plotted in atomic percentages, and the constant atomic ratio and electron · atom ratio lines are also shown.

Reference to figure 24 shows that, as the zinc percentage of the alloy from which the crystals were extracted increases, the composition of the primary T_1 crystals moves initially along the line *ab*, but for alloy zinc contents of more than 55 % of zinc, begins to deviate from this line. In this connexion, it should be remembered that the temperature at which the primary separation occurs decreases as the zinc content of the alloy increases. The points plotted therefore correspond with the compositions of the T_1 phase at different temperatures

(*c*) *The compound* $MnAl_4$

According to the results given in figure 23, the primary $MnAl_4$ field extends, at low manganese contents, only from approximately 69 to 76 % of zinc All the specimens analysed for both manganese and zinc, except that from alloy 1·25/75, therefore correspond with secondary crystals. Since the phase is clearly of variable composition, the compositions of secondary crystals represent an average only. It is this factor which accounts for the close agreement of the analysis from all samples except that from alloy 1·25/75 This extract, since it was obtained from an alloy at the extreme zinc-rich end of the region in which $MnAl_4$ crystallizes as the primary constituent, probably represents roughly the limiting solubility of zinc in $MnAl_4$ Zinc therefore dissolves in the compound $MnAl_4$ up to a limit of approximately 5·2 %, the substitution of zinc for manganese being approximately atom for atom.

(*d*) *The ternary compound* T_2

The composition of the ternary compound T_2 is also of interest, since it corresponds accurately with the formula Mn_2ZnAl_9. The electron : atom ratio of this compound, according to the scheme outlined above, is 1·81 If, however, the contribution to the structure of two atoms of manganese together with one of zinc is considered, it is seen that there remain 5·32 vacancies to be filled by electrons from the structure as a whole This is very similar to the contribution (5·13 vacancies) of three cobalt atoms, and it is possible that the compound is analogous to $CoAl_3$, which has an electron · atom ratio of 1·82.

(*e*) *The ternary compound* T_3

The ternary compound T_3 cannot be represented by as simple a formula as the compound T_2. It appears to conform most closely with the arrangement $(MnZn)_5Al_{11}$, or $Mn_{2\cdot2}Zn_{2\cdot2}Al_{11}$. It may be significant that the electron : atom ratio of this compound is close to that of $MnAl_4$, as shown in figure 24.

From the results of these experiments, therefore, it is clear that both size and electronic factors may operate together in intermetallic systems of the type examined.



FIGURE 8

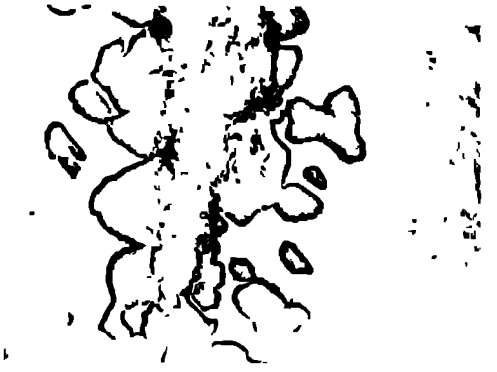


FIGURE 9



FIGURE 10



FIGURE 11

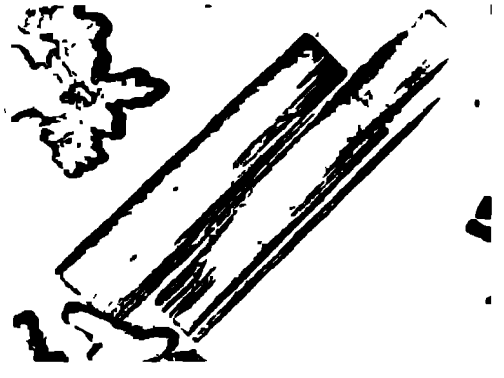


FIGURE 12

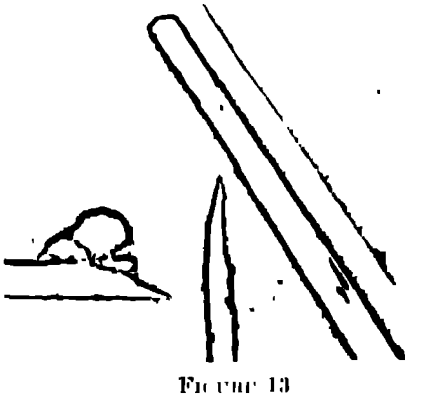


FIGURE 13

(*Facing p.* 100)



FIGURE 14



FIGURE 15



FIGURE 16



FIGURE 17



FIGURE 18

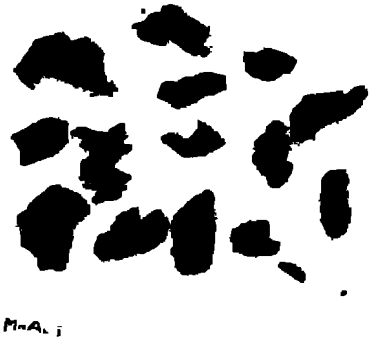


FIGURE 19

In particular, it is very probable that, when present as a minor constituent in alloys based upon a solvent of relatively high valency, transitional metals absorb electrons from the structure as a whole by filling up the vacancies in the atomic orbitals postulated by the Pauling theory.

This research was carried out in the Metallurgy Department of the University of Birmingham under the general supervision of Professor D. Hanson, D Sc , to whom the authors' thanks are due for his interest and support. The authors must express their gratitude to Mr G. Welsh for valuable assistance with some of the experimental work. Grateful acknowledgement is also made to the Royal Society, the Chemical Society, Imperial Chemical Industries, Ltd , and the Department of Scientific and Industrial Research for financial assistance and other facilities which have greatly encouraged the work

References

Butchers, E. & Hume-Rothery, W. 1945 *J. Inst. Met.* **71**, 87
Dix, E. H , Jr., Fink, W. L. & Willey, L. A. 1933 *Trans. Amer Inst Min. Met. Engrs*, **104**, 335
Gebhardt, E. 1942 *Z Metallk.* **34**, 259.
Köster, W. & Bechthold, W. 1938 *Z. Metallk* **30**, 294.
Pauling, L 1938 *Phys. Rev* (ii), **54**, 899.
Phillips, H W L. 1943 *J. Inst Met* **69**, 275.
Raynor, G. V. 1943 *Institute of Metals Annotated Equilibrium Diagram Series*, no 1 London.
Raynor, G. V 1944*a* *J Inst. Met.* **70**, 507
Raynor, G. V 1944*b* *J Inst Met* **70**, 531.
Raynor, G V. 1945 *Phil. Mag.* **36**, 770.
Raynor, G V. & Hume-Rothery, W 1943 *J Inst Met* **69**, 415.
Schramm, J. 1940 *Z Metallk* **32**, 399

Description of Plates 3 and 4

Plate 3

Figure 8. Crystals of $MnAl_6$ with spines of $MnAl_4$, alloy 3·0/30 × 125.

Figure 9. Cored crystal in alloy 3 0/70 × 125.

Figure 10 Dendritic T_3 crystals and plates of $MnAl_3$, alloy 2 0/80 × 125

Figure 11 Polygonal T_1 crystal in alloy 2 0/55. × 125.

Figure 12. Needles of $MnAl_4$ in alloy 1·25/75. × 125.

Figure 13 Cross-sections of plates of T_2 in alloy 1·25/75. × 125

Plate 4

Figure 14. Extracted crystals of $MnAl_6$. × 2·5.

Figure 15 Extracted crystals of ternary compound T_1. × 5

Figure 16 Extracted crystals of $MnAl_4$. × 5.

Figure 17. Extracted crystals of ternary compound T_2. × 5.

Figure 18. Extracted crystals of ternary compound T_3. × 5.

Figure 19. Extracted crystals of $MnAl_3$. × 5.

X-ray reflexions from dilute solid solutions

By K. Huang, *H. H. Wills Physical Laboratory, University of Bristol*

(*Communicated by N. F Mott, F.R S —Received 24 September* 1946)

The effect on X-ray reflexion of deviations of the atoms from the ideal lattice sites caused by the presence of randomly distributed foreign atoms in a dilute solid solution is investigated quantitatively The form of the function used to describe these deviations is suggested by the distortions produced in an elastic medium by a number of spherically symmetric point centres of distortion Hence one is led to two types of effects, exactly analogous to the thermal effects (i) a weakening of the ordinary interference maxima, (ii) the presence of 'diffuse maxima' associated with the ordinary maxima The change of lattice constant appears naturally in the analysis It is used to determine the magnitude of the effects (i) and (ii) By applying the theoretical formulae to the solid solutions Au-Cu, we find that it should be possible to detect (i) experimentally, the thermal effect is secondary and cannot mask the distortion effect But the effect (ii) mixes with the thermal diffuse maxima and is found to be very much smaller at ordinary temperatures This conclusion is, however, not regarded as general, especially in view of the anisotropic nature of the thermal effect.

When the foreign atoms in a dilute solid solution are distributed randomly among the lattice sites, X-ray reflexion from the crystal should be modified as compared with pure specimens, for two principal reasons In the first place, the intrinsic scattering power of the foreign atoms is different This effect was treated generally by von Laue (1918) many years ago. In the second place, when the foreign atoms are of a different atomic radius, e.g Ag, Au in Cu or vice versa, all the atoms in the crystal are pulled slightly out of the regular sites appropriate to the lattice type. These random deviations from the regular lattice should modify X-ray reflexions in the same way as thermal agitation. In the present paper, a quantitative treatment of the second effect is attempted, it being based on a simple assumption about the distortion round a dissolved atom It will be seen that Laue's results on the first effect can be simply combined with the result given here

Considering a crystal lattice formed of randomly distributed atoms of two kinds mixed in comparable proportions, it must obviously be extremely difficult to describe quantitatively the distorted configuration But when the solution is dilute, a very natural suggestion is to regard the lattice as an elastic medium with centres of distortion at the sites of the foreign atoms When the state of distortion around the centres is regarded as largely independent of the effects of other centres, the distortion of the medium can be described by a simple superposition of the effects of the individual centres Thus if $\mathbf{u}(\mathbf{r})$ approximately describe the elastic displacement at a point $\mathbf{r}$ under the influence of a single foreign atom at the origin, the displacement of an atom at a lattice site $\mathbf{a}_n$ in a lattice, where a number of sites are occupied by foreign atoms, can be described approximately by

$$\sum_{i=\text{occ.}} \mathbf{u}(\mathbf{a}_n - \mathbf{a}_i), \tag{1}$$

where i is summed over the sites occupied by the foreign atoms.

In estimating the distortion the anisotropy of the medium will be neglected This leads to a unique choice of $\mathbf{u}(\mathbf{r})$, for the only suitable radially symmetric solution of the equation of elasticity in an isotropic medium is

$$\mathbf{u}(\mathbf{r}) = \frac{c\mathbf{r}}{|\mathbf{r}|^3}, \tag{2}$$

where c is a constant. The very general manner in which (1) and (2) have been arrived at suggests that they should reproduce the essential features of the distortion without too much inaccuracy.

Consider a crystal of such small size that the extinction of the incident X-rays within it can be neglected If the crystal lies in an incident beam of wave-length λ and direction $\mathbf{n}_0$, the intensity of the reflected X-ray in the direction $\mathbf{n}$ is given by

$$Sf^2 \left| \sum_n e^{i(\xi \cdot \mathbf{r}_n)} \right|^2, \tag{3}$$

where $\xi = (2\pi/\lambda)(\mathbf{n} - \mathbf{n}_0)$, $\mathbf{n}$ and $\mathbf{n}_0$ are unit vectors along the reflected and incident directions respectively, $\mathbf{r}_n$ is the position vector of the atom n in the lattice and the summation extends over all atoms Ignore, for the moment, the different scattering power of the foreign atoms So f stands for the atomic scattering factor of the solvent atoms and S is the intensity scattered by a single electron as given by the formula of J. J Thomson.

Using (1), $\mathbf{r}_n$ can be written explicitly as

$$\mathbf{r}_n = \mathbf{a}_n + \sum_i \mathbf{u}(\mathbf{a}_n - \mathbf{a}_i).$$

When this is substituted in (3), and the square is multiplied out, (3) becomes

$$Sf^2 \sum_n \sum_m \exp\{i\xi \cdot [\mathbf{a}_n - \mathbf{a}_m + \sum_i (\mathbf{u}(\mathbf{a}_n - \mathbf{a}_i) - \mathbf{u}(\mathbf{a}_m - \mathbf{a}_i))]\}. \tag{4}$$

For almost all pairs of $\mathbf{a}_n$, $\mathbf{a}_m$, there are a great number of terms corresponding to the same value of $\mathbf{a}_n - \mathbf{a}_m$. These related terms are different from one another, because the distribution of foreign atoms characterized by the indices appearing in the summation over i is different around each pair of atoms $\mathbf{a}_n$, $\mathbf{a}_m$ Since each group of related terms is summed over a large number of pairs of atoms, a statistical mean value can be substituted for each of the terms This mean value is the statistical average of the term belonging to any pair $\mathbf{a}_n$, $\mathbf{a}_m$ of the group for all possible distributions of the foreign atoms over the lattice An average taken in this sense will be indicated by a horizontal bar over the expression concerned.

Rewrite (4) as

$$Sf^2 \sum_n \sum_m \exp\{i\xi \cdot [\mathbf{a}_n - \mathbf{a}_m + \sum_i \overline{(\mathbf{u}(\mathbf{a}_n - \mathbf{a}_i) - \mathbf{u}(\mathbf{a}_m - \mathbf{a}_i))}]\}$$
$$\times \exp\{i\xi \cdot [\sum_i (\mathbf{u}(\mathbf{a}_n - \mathbf{a}_i) - \mathbf{u}(\mathbf{a}_m - \mathbf{a}_i)) - \sum_i \overline{(\mathbf{u}(\mathbf{a}_n - \mathbf{a}_i) - \mathbf{u}(\mathbf{a}_m - \mathbf{a}_i))}]\}. \tag{5}$$

Now the expression in the square bracket in the second factor is of the nature of a fluctuation from the mean value, so its mean value can be calculated approximately

by expanding the exponential function As the first factor is not affected in taking the average, at present the second factor will be considered alone.

Expanding and retaining the first three terms, the mean value of the second factor is obtained·

$$1-\tfrac{1}{2}\overline{\{\xi\ [\sum_i(\mathbf{u}(\mathbf{a}_n-\mathbf{a}_i)-\mathbf{u}(\mathbf{a}_m-\mathbf{a}_i))-\sum_i(\overline{\mathbf{u}(\mathbf{a}_n-\mathbf{a}_i)-\mathbf{u}(\mathbf{a}_m-\mathbf{a}_i)})]\}^2},$$

which reduces to

$$1-\tfrac{1}{2}\overline{\{\xi\ \sum_i(\mathbf{u}(\mathbf{a}_n-\mathbf{a}_i)-\mathbf{u}(\mathbf{a}_m-\mathbf{a}_i))\}^2}+\tfrac{1}{2}\{\overline{\xi\ \sum_i(\mathbf{u}(\mathbf{a}_n-\mathbf{a}_i)-\mathbf{u}(\mathbf{a}_m-\mathbf{a}_i))}\}^2,$$

and hence to

$$\begin{aligned}1-\tfrac{1}{2}\{&\overline{\sum_i[\xi\ (\mathbf{u}(\mathbf{a}_n-\mathbf{a}_i)-\mathbf{u}(\mathbf{a}_m-\mathbf{a}_i))]^2}\\&+\overline{\sum_{i\neq j}[\xi\cdot(\mathbf{u}(\mathbf{a}_n-\mathbf{a}_i)-\mathbf{u}(\mathbf{a}_m-\mathbf{a}_i))]\,[\xi\ (\mathbf{u}(\mathbf{a}_n-\mathbf{a}_j)-\mathbf{u}(\mathbf{a}_m-\mathbf{a}_j))]}\}\\&+\tfrac{1}{2}\{\xi\cdot\overline{\sum_i(\mathbf{u}(\mathbf{a}_n-\mathbf{a}_i)-\mathbf{u}(\mathbf{a}_m-\mathbf{a}_i))}\}^2\end{aligned}\qquad(6)$$

In working out the mean values of the summations in the first bracket, add the summations for all possible distributions of i over the lattice sites and then divide by the number of such distributions. For the first summation in the first bracket, attention is directed to a particular lattice site The number of times this site appears in the summation over i divided by the total number of times the summation is repeated (for various distributions) is just the atomic concentration p, of the foreign atoms It follows that the mean value is just p times the summation extended over all sites (except of course $\mathbf{a}_n$ and $\mathbf{a}_m$, at which $\mathbf{u}$ has no meaning) To a good approximation, the probabilities of occupation of two sites i and j can be regarded as entirely independent Then it follows similarly that the mean value of the second summation is given by p^2 times the summation with i, j extended over all lattice sites

The mean value in the last bracket of (6) may be evaluated similarly and (6) can be written as

$$\begin{aligned}1-\tfrac{1}{2}\{&p\sum_i^{\text{all sites}}[\xi\cdot(\mathbf{u}(\mathbf{a}_n-\mathbf{a}_i)-\mathbf{u}(\mathbf{a}_m-\mathbf{a}_i))]^2\\&+p^2\sum_{i\neq j}^{\text{all sites}}[\xi\cdot(\mathbf{u}(\mathbf{a}_n-\mathbf{a}_i)-\mathbf{u}(\mathbf{a}_m-\mathbf{a}_i))]\,[\xi\cdot(\mathbf{u}(\mathbf{a}_n-\mathbf{a}_j)-\mathbf{u}(\mathbf{a}_m-\mathbf{a}_j))]\}\\&+\tfrac{1}{2}p^2\{\sum_i^{\text{all sites}}\xi\ (\mathbf{u}(\mathbf{a}_n-\mathbf{a}_i)-\mathbf{u}(\mathbf{a}_m-\mathbf{a}_i))\}^2,\end{aligned}$$

which reduces to

$$1-\tfrac{1}{2}(p-p^2)\sum_i^{\text{all sites}}[\xi\cdot(\mathbf{u}(\mathbf{a}_n-\mathbf{a}_i)-\mathbf{u}(\mathbf{a}_m-\mathbf{a}_i))]^2.$$

Substituting this expression in (5), then, for the reflected intensity,

$$Sf^2 \sum_n \sum_m \exp\{i\boldsymbol{\xi}\,[\mathbf{a}_n-\mathbf{a}_m+p\sum_t^{\text{all sites}}(\mathbf{u}(\mathbf{a}_n-\mathbf{a}_t)-\mathbf{u}(\mathbf{a}_m-\mathbf{a}_t))]\}$$

$$\times\{1-\tfrac{1}{2}(1-\delta_{nm})(p-p^2)\sum_t^{\text{all sites}}[\boldsymbol{\xi}.(\mathbf{u}(\mathbf{a}_n-\mathbf{a}_t)-\mathbf{u}(\mathbf{a}_m-\mathbf{a}_t))]^2\}, \qquad (7)$$

where the mean value in the exponential function is transformed into the ordinary summation as the last term in (6).

When the explicit form (2) for $\mathbf{u}$ is introduced, it is seen that the following three summations are involved

$$\text{(I)}\quad c\sum_t \boldsymbol{\xi}.\left(\frac{\mathbf{a}_n-\mathbf{a}_t}{|\mathbf{a}_n-\mathbf{a}_t|^3}-\frac{\mathbf{a}_m-\mathbf{a}_t}{|\mathbf{a}_m-\mathbf{a}_t|^3}\right),$$

$$\text{(II)}\quad c^2\sum_t \frac{[\boldsymbol{\xi}.(\mathbf{a}_n-\mathbf{a}_t)]^2}{|\mathbf{a}_n-\mathbf{a}_t|^6},$$

$$\text{(III)}\quad c^2\sum_t \frac{[\boldsymbol{\xi}.(\mathbf{a}_n-\mathbf{a}_t)][\boldsymbol{\xi}.(\mathbf{a}_m-\mathbf{a}_t)]}{|\mathbf{a}_n-\mathbf{a}_t|^3\,|\mathbf{a}_m-\mathbf{a}_t|^3}.$$

(II) and (III) together form the term in the second bracket of (7) All three will be considered in turn

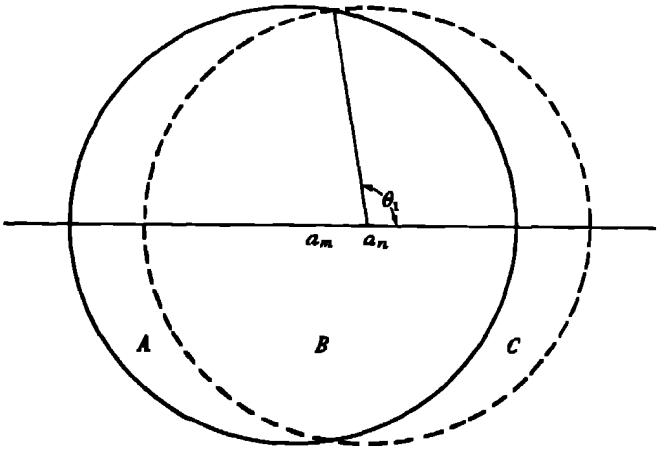


FIGURE 1

Consider (I) first To work out such a lattice sum accurately, the summations must be carried out within a certain distance about the singular points, but may be replaced by integrals beyond. Draw a sphere about $\mathbf{a}_m$ with radius R, large compared with both this distance and with $|\mathbf{a}_n-\mathbf{a}_m|$, as shown in the figure by the sphere in solid line. To obtain the contributions to (I) of lattice points within the sphere, the following simple device can be used to avoid the complexities near $\mathbf{a}_n$ and $\mathbf{a}_m$. Another sphere of equal radius about $\mathbf{a}_n$ can be drawn as shown in figure 1 in dotted line. The three regions A, B and C can be distinguished, which are

respectively the spaces included by first sphere alone, by both spheres and by second sphere alone Correspondingly

$$S_A(R)=\sum_i^A \frac{\xi\cdot(\mathbf{a}_m-\mathbf{a}_i)}{|\mathbf{a}_m-\mathbf{a}_i|^3},\quad S_B(R)=\sum_i^B \frac{\xi\cdot(\mathbf{a}_m-\mathbf{a}_i)}{|\mathbf{a}_m-\mathbf{a}_i|^3},\quad S_C(R)=\sum_i^C \frac{\xi\cdot(\mathbf{a}_m-\mathbf{a}_i)}{|\mathbf{a}_m-\mathbf{a}_i|^3},$$

and $$S'_A(R)=\sum_i^A \frac{\xi\cdot(\mathbf{a}_n-\mathbf{a}_i)}{|\mathbf{a}_n-\mathbf{a}_i|^3},\quad S'_B(R)=\sum_i^B \frac{\xi\cdot(\mathbf{a}_n-\mathbf{a}_i)}{|\mathbf{a}_n-\mathbf{a}_i|^3},\quad S'_C(R)=\sum_i^C \frac{\xi\cdot(\mathbf{a}_n-\mathbf{a}_i)}{|\mathbf{a}_n-\mathbf{a}_i|^3},$$

where i is to be summed over the lattice sites within the region indicated

The contribution to (I) of lattice sites within the solid sphere is

$$S'_A(R)-S_A(R)+S'_B(R)-S_B(R) \tag{8}$$

Now $$S_A(R)+S_B(R)=S'_B(R)+S'_C(R)$$

or $$S_A(R)+S_B(R)-S'_B(R)-S'_C(R)=0$$

Adding this expression to (8), one obtains for this contribution

$$S'_A(R)-S'_C(R).$$

Within regions A and C, the sums may be replaced by integrals Denote, for brevity, $|\mathbf{a}_n-\mathbf{a}_m|$ by a and the angle between $\mathbf{a}_n-\mathbf{a}_m$ and ξ by Θ, and use spherical polar co-ordinates (r,θ,ϕ) about the point $\mathbf{a}_n$, with polar axis along $\mathbf{a}_n-\mathbf{a}_m$, then

$$\begin{aligned}S'_A(R)-S'_C(R)&=\frac{|\xi|}{v}\left\{\int_0^{2\pi}d\phi\int_0^{\theta_1}\sin\theta\,d\theta\int_{-a\cos\theta+\sqrt{(R^2-a^2\sin^2\theta)}}^{R}(\cos\Theta\cos\theta+\sin\Theta\sin\theta\cos\phi)\,dr\right.\\&\left.-\int_0^{2\pi}d\phi\int_{\theta_1}^{\pi}\sin\theta\,d\theta\int_{R}^{-a\cos\theta+\sqrt{(R^2-a^2\cos^2\theta)}}(\cos\Theta\cos\theta+\sin\Theta\sin\theta\cos\phi)\,dr\right\},\end{aligned}$$

where θ_1 is the angle indicated in the figure, and v is the atomic volume. This leads directly to

$$\begin{aligned}S'_A(R)-S'_C(R)&=\frac{2\pi|\xi|\cos\Theta}{v}\int_0^{\pi}[R+a\cos\theta-\sqrt{(R^2-a^2\sin^2\theta)}]\cos\theta\sin\theta\,d\theta\\&\simeq\frac{4\pi\xi\cdot(\mathbf{a}_n-\mathbf{a}_m)}{3v}\end{aligned}$$

for large R. Therefore neglecting boundary effects one finds for (I)

$$\text{(I)}\quad c\sum_i\xi\left(\frac{\mathbf{a}_n-\mathbf{a}_i}{|\mathbf{a}_n-\mathbf{a}_i|^3}-\frac{\mathbf{a}_m-\mathbf{a}_i}{|\mathbf{a}_m-\mathbf{a}_i|^3}\right)=\frac{4\pi c\xi\cdot(\mathbf{a}_n-\mathbf{a}_m)}{3v}. \tag{9}$$

This simply represents a uniform linear expansion with the strain $4\pi c/3v$. The same result may also be obtained by general arguments on volume expansion, but the above deduction is preferable, because of its greater clarity.

(II) is the simplest of the three sums and occurs in problems of lattice dynamics. Its rigorous numerical values have been given for some simple lattices in works by

Born (1940) on lattice stability. But it will be found useful in connexion with the evaluation of (III) to work out (II) by a crude method. This consists in replacing the summation by an integral throughout, i e.

$$\text{(II)} \quad c^2 \sum_i \frac{[\boldsymbol{\xi}.(\mathbf{a}_n-\mathbf{a}_i)]^2}{|\mathbf{a}_n-\mathbf{a}_i|^6} \simeq \frac{c^2|\boldsymbol{\xi}|^2}{v}\int_{\rho_0<r<\infty}\frac{z^2}{r^6}d\tau = \frac{4\pi c^2|\boldsymbol{\xi}|^2}{3\rho_0 v}. \tag{10}$$

ρ_0 cannot be chosen exactly but is of the order of the interatomic distance

(III) is considerably more complicated To evaluate it generally, one resorts to the crude method For brevity, call the angle between $\boldsymbol{\xi}$ and $\mathbf{a}_n-\mathbf{a}_m$, Θ, but this time use $2a$ for $|\mathbf{a}_n-\mathbf{a}_m|$ Write the integral approximation first in co-ordinates (x, y, ϕ), where (x, y) are Cartesian co-ordinates in a plane passing through both $\mathbf{a}_n$ and $\mathbf{a}_m$ with X-axis from $\mathbf{a}_m$ to $\mathbf{a}_n$, and ϕ is the azimuth angle of the plane as measured from the plane that is parallel to $\boldsymbol{\xi}$

$$c^2\sum_i \frac{[\boldsymbol{\xi}\ (\mathbf{a}_n-\mathbf{a}_i)][\boldsymbol{\xi}.(\mathbf{a}_m-\mathbf{a}_i)]}{|\boldsymbol{\xi}\ (\mathbf{a}_n-\mathbf{a}_i)|^3|\boldsymbol{\xi}\ (\mathbf{a}_m-\mathbf{a}_i)|^3}$$

$$\simeq \frac{c^2|\boldsymbol{\xi}|^2}{v}\int_0^{2\pi}d\phi\iint_{(r_1,r_2>\rho_0)} y\,dy\,dx$$

$$\times \frac{(\cos\Theta\cos\theta_1+\sin\Theta\sin\theta_1\cos\phi)(-\cos\Theta\cos\theta_2+\sin\Theta\sin\theta_2\cos\phi)}{r_1^2r_2^2}$$

$$\simeq \frac{\pi c^2|\boldsymbol{\xi}|^2}{v}\iint_{(r_1,r_2>\rho_0)} y\,dy\,dx\,\frac{(-2\cos^2\Theta\cos\theta_1\cos\theta_2+\sin^2\Theta\sin\theta_1\sin\theta_2)}{r_1^2r_2^2},$$

$r_1, r_2, \theta_1, \theta_2$ are the quantities as indicated in figure 2.

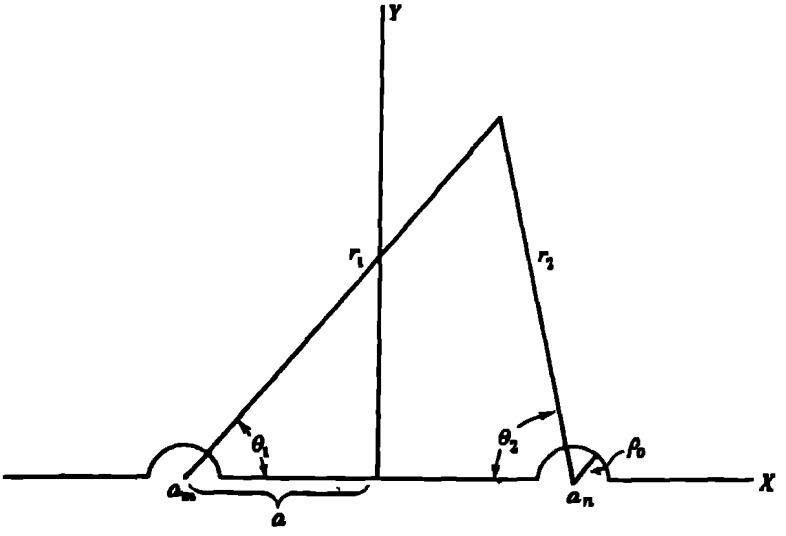


FIGURE 2

It is to be noticed that only the half-plane $y > 0$ is involved in the Cartesian co-ordinates used here. Furthermore, as the integrand is symmetric with respect to the y-axis, the integration has only to be carried through the first quadrant and the result doubled. The simplest way to effect this integration is by introducing the plane elliptic co-ordinates

$$x = \frac{\mu\lambda}{a}, \quad y = \frac{1}{a}\sqrt{[(a^2-\mu^2)(\lambda^2-a^2)]}$$

for the first quadrant $\qquad 0 \leqslant \mu \leqslant a, \quad a \leqslant \lambda \leqslant \infty$

In terms of these co-ordinates

$$r_1 = \lambda - \mu, \quad r_2 = \lambda + \mu;$$

$$y\frac{\partial(x, y)}{\partial(\lambda, \mu)} = \frac{\partial(x, y^2)}{\partial(\lambda, \mu)} = \frac{(\lambda^2-\mu^2)}{a}.$$

Transformation of the integrand is simple, if the trigonometric functions are first expressed in terms of r_1 and r_2. Remembering that the equation of the semicircular boundary is now simply given by

$$\lambda - \mu = \rho_0,$$

the integral is obtained in terms of λ, μ:

$$\frac{2\pi c^2 |\xi|^2}{a^3 v}\left\{2\cos^2\Theta\int_a^{a+\rho_0} d\lambda\int_0^{\lambda-\rho_0} d\mu\,\frac{(\lambda^2\mu^2-a^4)}{(\lambda^2-\mu^2)^2} + 2\cos^2\Theta\int_{a+\rho_0}^{\infty} d\lambda\int_0^{a} d\mu\,\frac{(\lambda^2\mu^2-a^4)}{(\lambda^2-\mu^2)^2}\right.$$
$$+\sin^2\Theta\int_a^{a+\rho_0} d\lambda\int_0^{\lambda-\rho_0} d\mu\,\frac{a^2(\lambda^2+\mu^2)-\lambda^2\mu^2-a^4}{(\lambda^2-\mu^2)^2}$$
$$\left.+\sin^2\Theta\int_{a+\rho_0}^{\infty} d\lambda\int_0^{a} d\mu\,\frac{a^2(\lambda^2+\mu^2)-\lambda^2\mu^2-a^4}{(\lambda^2-\mu^2)^2}\right\}$$

The integration is tedious but presents no difficulty.

Retaining Θ as the angle between ξ and $\mathbf{a}_n - \mathbf{a}_m$, but replacing $2a$ by $\mathbf{a}_n - \mathbf{a}_m$, the approximate result of (III) can be written

$$\text{(III)} \quad c^2\sum_i \frac{[\xi\cdot(\mathbf{a}_n-\mathbf{a}_i)][\xi\cdot(\mathbf{a}_m-\mathbf{a}_i)]}{|\mathbf{a}_n-\mathbf{a}_i|^3\,|\mathbf{a}_m-\mathbf{a}_i|^3}$$
$$= \frac{\pi c^2|\xi|^2}{v}\left\{\frac{2\sin^2\Theta}{|\mathbf{a}_n-\mathbf{a}_m|} - \frac{(2\rho_0)^2\sin^2\Theta}{3|\mathbf{a}_n-\mathbf{a}_m|^3} + \frac{2(2\rho_0)^2\cos^2\Theta}{3|\mathbf{a}_n-\mathbf{a}_m|^3}\right\}. \qquad (11)$$

Substituting the values of (I), (II) and (III) as given by (9), (10) and (11) in the intensity formula (7), one can write in place of (7)

$$Sf^2\sum_n\sum_m\left[1-(1-\delta_{nm})(p-p^2)\frac{4\pi c^2|\xi|^2}{3\rho_0 v}\right]e^{i\xi\cdot(\mathbf{a}_n-\mathbf{a}_m)(1+(4\pi cp)/(3v))}$$
$$+Sf^2(p-p^2)\frac{\pi c^2|\xi|^2}{2v}\sum_n\sum_m(1-\delta_{nm})\left\{\frac{2\sin^2\Theta}{|\mathbf{a}_n-\mathbf{a}_m|} - \frac{(2\rho_0)^2\sin^2\Theta}{3|\mathbf{a}_n-\mathbf{a}_m|^3} + \frac{2(2\rho_0)^2\cos^2\Theta}{3|\mathbf{a}_n-\mathbf{a}_m|^3}\right\}$$
$$\times e^{i\xi\cdot(\mathbf{a}_n-\mathbf{a}_m)(1+(4\pi cp)/(3v))} \qquad (12)$$

The first part is exactly analogous to the kind of thermal effect discussed by Debye (1914) It results from independent displacements of separate atoms The sharpness of the pattern is not affected by it, only the intensity of the maxima is reduced, and a diffuse background of slowly varying intensity is produced. In the present case, the intensity of the maximum at ξ_0 is reduced by the factor $(p-p^2)\,4\pi c^2\,|\,\xi_0\,|^2/(3\rho_0 v)$. The diffuse background intensity is given by $NSf^2(p-p^2)\{4\pi c^2\,|\,\xi_0\,|^2/(3\rho_0 v)\}$, where N is the total number of atoms in the crystal A more careful discussion of this effect will be discussed later

There is also the additional factor $1+(4\pi cp)/(3v)$ in the exponential This gives the effect of a uniform expansion of the lattice by this factor in every direction—an effect well verified by experiments

The term in the second line results from the fact that the relative displacement of atoms originally at $\mathbf{a}_n$ and $\mathbf{a}_m$ has an appreciable dependence upon $\mathbf{a}_n-\mathbf{a}_m$, when they are near. The first line is the effect of a constant mean fluctuation for all groups of $\mathbf{a}_n-\mathbf{a}_m$, equal to that for large separation. The second line comes in as a correction for rather closely situated pairs Since the recently discovered diffuse maxima (Lonsdale 1942) due to thermally excited elastic waves arise from a cause of an entirely similar nature, it would be expected that the second line of (12) would also lead to such diffuse maxima

For convenience, denote the first part of the intensity (12) by J_1, the rest by J_2 Proceed to reduce J_2 to a form that can be interpreted Since J_2 itself is the small effect to be investigated, the small correction $(4\pi cp)/(3v)$ in the exponential may be omitted The sum will be evaluated by using the Fourier theorem. For this purpose, it is necessary to find the Fourier transform $F(x,y,z)=F(\mathbf{r})$ of the expression included in the curved bracket of J_2 in (12) This expression actually holds only for $|\,\mathbf{a}_n-\mathbf{a}_m\,|>2\rho_0$ To give the expression a more definite meaning as a function of $|\,\mathbf{a}_n-\mathbf{a}_m\,|$, it will be cut off rather arbitrarily at $|\,\mathbf{a}_n-\mathbf{a}_m\,|=2\rho_0$, i e the expression will be regarded as vanishing for $|\,\mathbf{a}_n-\mathbf{a}_m\,|<2\rho_0$ Then if (x,y,z) are co-ordinates of the position vector in any definite Cartesian co-ordinate system in space and (λ,μ,ν) the corresponding components of $\mathbf{a}_n-\mathbf{a}_m$, the Fourier transform is defined as

$$F(x,y,z)=\frac{1}{(2\pi)^{\frac{3}{2}}}\int d\lambda\int d\mu\int d\nu_{(\lambda^2+\mu^2+\nu^2>4\rho_0^2)}\times\left\{\frac{2\sin^2\Theta}{(\lambda^2+\mu^2+\nu^2)^{\frac{1}{2}}}-\frac{(2\rho_0)^2\sin^2\Theta}{3(\lambda^2+\mu^2+\nu^2)^{\frac{1}{2}}}+\frac{2(2\rho_0)^2\cos^2\Theta}{3(\lambda^2+\mu^2+\nu^2)^{\frac{1}{2}}}\right\}e^{-i(\lambda x+\mu y+\nu z)},\qquad(13)$$

where Θ is the angle between ξ and (λ,μ,ν) For the integration in (13), ξ and $\mathbf{r}(x,y,z)$ are two fixed vectors in space, so the polar co-ordinates (ρ,θ,ϕ) may be introduced in place of the Cartesian co-ordinates (λ,μ,ν), with polar axis along $\mathbf{r}(x,y,z)$ and ξ in the plane $\phi=0$. If the angle between ξ and $\mathbf{r}$ be called α, then (13) becomes

$$F(x,y,z)=\frac{1}{(2\pi)^{\frac{3}{2}}}\int_0^{\pi}\sin\theta\,d\theta\int_{2\rho_0}^{R\to\infty}\rho^2 d\rho\left\{\frac{1}{\rho}[(2-\sin^2\alpha)-(2\cos^2\alpha-\sin^2\alpha)\cos^2\theta]+\frac{(2\rho_0)^2}{3\rho^3}[(-2+3\sin^2\alpha)+(6\cos^2\alpha-\sin^2\alpha)\cos^2\theta]\right\}e^{-i|\mathbf{r}|\rho\cos\theta},$$

where, for shortness, the simple integration over ϕ has been performed, and R is eventually to tend to ∞ as indicated. The subsequent integrations are simple and lead to

$$F(x,y,z) = \lim_{R\to\infty}\left(\frac{2}{\pi}\right)^{\frac{1}{2}}\frac{1}{r^2}\left\{2\left[(1-\cos^2\alpha)(\cos 2r\rho_0 - \cos Rr)\right.\right.$$
$$+(2\cos^2\alpha-\sin^2\alpha)\left(\frac{\sin 2r\rho_0}{2r\rho_0}-\frac{\sin Rr}{Rr}\right)\Big]$$
$$\left.+(\tfrac{1}{3}-\cos^2\alpha)\left[\frac{\sin 2\rho_0 r}{2\rho_0 r}-\cos 2\rho_0 r-\left(\frac{2\rho_0}{R}\right)^2\left(\frac{\sin Rr}{Rr}-\cos Rr\right)\right]\right\}$$
$$=\left(\frac{2}{\pi}\right)^{\frac{1}{2}}\frac{1}{r^2}\left\{(5\cos^2\alpha-\tfrac{5}{3})\frac{\sin 2\rho_0 r}{2\rho_0 r}+(\tfrac{5}{3}-\cos^2\alpha)\cos 2\rho_0 r\right\}. \tag{14}$$

$\lim_{R\to\infty}\cos Rr$ has been put equal to zero, for $F(x,y,z)$ will be used eventually in an integration over (x,y,z), as $R\to\infty$, $\cos Rr$ oscillates infinitely rapidly so contributes nothing on integration. For simplicity r is used for $|r| = \sqrt{(x^2+y^2+z^2)}$ in the formula.

Comparing the definition of $F(x,y,z)$ as given by (13) with (12), then by applying directly Fourier's theorem

$$J_2 = Sf^2(p-p^2)\frac{\pi c^2|\xi|^2}{(2\pi)^{\frac{1}{2}}v}\int dx\int dy\int dz\, F(x,y,z)\,|\sum_n e^{i(\xi+\tau)\cdot\mathbf{a}_n}|^2.$$

The factor $(1-\delta_{nm})$ has made no difference, because the expression in the curved bracket of (12) has been assumed to be zero for $|\mathbf{a}_n-\mathbf{a}_m|<2\rho_0$ in obtaining $F(xyz)$.

To work out the summation over $\mathbf{a}_n$, $\mathbf{a}_m$ explicitly, the type of lattice must be specified, which has been left undecided so far With the simple alloy systems formed of Cu, Ag, Au in mind, the expression for J_2 will be worked out for the case of a face-centred cubic lattice with lattice constant d Splitting up the summation into two factors, one a summation over the four sites in the cubic unit cell, the other a summation over the cells and simplifying in the usual way, then

$$J_2 = Sf^2(p-p^2)\frac{\pi c^2|\xi|^2}{(2\pi)^{\frac{1}{2}}v}\int dx\int dy\int dz\, F(x,y,z)$$
$$\times\left|1+e^{i\frac{1}{2}(\xi_x+x)d+i\frac{1}{2}(\xi_y+y)d}+e^{i\frac{1}{2}(\xi_y+y)d+i\frac{1}{2}(\xi_z+z)d}+e^{i\frac{1}{2}(\xi_z+z)d+i\frac{1}{2}(\xi_x+x)d}\right|^2$$
$$\times\frac{\sin^2[\frac{1}{2}N^{\frac{1}{3}}(\xi_x+x)d]\sin^2[\frac{1}{2}N^{\frac{1}{3}}(\xi_y+y)d]\sin^2[\frac{1}{2}N^{\frac{1}{3}}(\xi_z+z)d]}{\sin^2[\frac{1}{2}(\xi_x+x)d]\sin^2[\frac{1}{2}(\xi_y+y)d]\sin^2[\frac{1}{2}(\xi_z+z)d]},$$

if the crystal is supposed to contain N atoms in the form of a cubic block. The last factor is well known to have vanishing value except in the immediate neighbourhood of

$$\xi_x+x=\frac{2\pi l}{d},\quad \xi_y+y=\frac{2\pi m}{d},\quad \xi_z+z=\frac{2\pi n}{d},\quad l,m,n=\text{integers},$$

where it rises to strong maxima. When r is not too near the origin, $F(x,y,z)$ behaves regularly So if ξ is not too near the points $(2\pi/d)(l,m,n)$, the first two factors may

be taken outside the integral signs at these points where the last factor attains its maxima. Therefore

$$J_2 = Sf^2(p-p^2)\frac{\pi c^2 |\xi|^2}{(2\pi)^4 v}\sum_{l,m,n} |1+e^{i\pi(l+m)}+e^{i\pi(m+n)}+e^{i\pi(n+l)}|^2 \times F\left(\frac{2\pi l}{d}-\xi_x, \frac{2\pi m}{d}-\xi_y, \frac{2\pi n}{d}-\xi_z\right) \times \int dx\int dy\int dz_{\text{near origin}}\left\{\frac{\sin^2(\frac{1}{2}N^{\frac{1}{3}}xd)\sin^2(\frac{1}{2}N^{\frac{1}{3}}yd)\sin^2(\frac{1}{2}N^{\frac{1}{3}}zd)}{\sin^2(\frac{1}{2}xd)\sin^2(\frac{1}{2}yd)\sin^2(\frac{1}{2}zd)}\right\}.$$

The integration can be carried out by using the approximate relations

$$\sin\tfrac{1}{2}xd \simeq \tfrac{1}{2}xd, \quad \sin\tfrac{1}{2}yd \simeq \tfrac{1}{2}yd, \quad \sin\tfrac{1}{2}zd \simeq \tfrac{1}{2}zd.$$

The complete result written out explicitly with the help of (14) becomes then

$$J_2 = NSf^2(p-p^2)\frac{\pi^2 c^2 |\xi|^2}{2d^3 v}\sum_{l,m,n} |1+e^{i\pi(l+m)}+e^{i\pi(m+n)}+e^{i\pi(n+l)}|^2 \times \frac{1}{\left(\frac{2\pi l}{d}-\xi_x\right)^2+\left(\frac{2\pi m}{d}-\xi_y\right)^2+\left(\frac{2\pi n}{d}-\xi_z\right)^2} \times \left\{\left(\cos 2\rho_0\left[\left(\frac{2\pi l}{d}-\xi_x\right)^2+\left(\frac{2\pi m}{d}-\xi_y\right)^2+\left(\frac{2\pi n}{d}-\xi_z\right)^2\right]^{\frac{1}{2}}\right) \times \left(\frac{5}{3}-\frac{\left[\xi_x\left(\frac{2\pi l}{d}-\xi_x\right)+\xi_y\left(\frac{2\pi m}{d}-\xi_y\right)+\xi_z\left(\frac{2\pi n}{d}-\xi_z\right)\right]^2}{[\xi_x^2+\xi_y^2+\xi_z^2]\left[\left(\frac{2\pi l}{d}-\xi_x\right)^2+\left(\frac{2\pi m}{d}-\xi_y\right)^2+\left(\frac{2\pi n}{d}-\xi_z\right)^2\right]}\right) + \frac{\sin 2\rho_0\left[\left(\frac{2\pi l}{d}-\xi_x\right)^2+\left(\frac{2\pi m}{d}-\xi_y\right)^2+\left(\frac{2\pi n}{d}-\xi_z\right)^2\right]^{\frac{1}{2}}}{2\rho_0\left[\left(\frac{2\pi l}{d}-\xi_x\right)^2+\left(\frac{2\pi n}{d}-\xi_y\right)^2+\left(\frac{2\pi n}{d}-\xi_z\right)^2\right]^{\frac{1}{2}}} \times \left(5\frac{\left[\xi_x\left(\frac{2\pi l}{d}-\xi_x\right)+\xi_y\left(\frac{2\pi m}{d}-\xi_y\right)+\xi_z\left(\frac{2\pi n}{d}-\xi_z\right)\right]^2}{[\xi_x^2+\xi_y^2+\xi_z^2]\left[\left(\frac{2\pi l}{d}-\xi_x\right)^2+\left(\frac{2\pi m}{d}-\xi_y\right)^2+\left(\frac{2\pi n}{d}-\xi_z\right)^2\right]}-\frac{5}{3}\right)\right\} \quad (15)$$

The first factor after the summation sign is just the structure factor of the unit cell of a face-centred cubic lattice. It is easy to see that it is 16, when l, m, n are either all even or all odd and zero otherwise. So the points $(2\pi/d)(l, m, n)$ involved actually in the summation form a body-centred cubic lattice with length of unit-cubic edge $= 4\pi/d$. Since the quantities actually involved in the summation are only distances of ξ from these points and angles between ξ and the lines joining it to these points, if one draws the equivalent polyhedra formed of faces that are perpendicular bisecting planes of first and second nearest neighbours in the lattice of the points $(2\pi/d)(l, m, n)$, J_2 depends only on the direction of ξ and its position in the polyhedron

that encloses ξ. Call the centre of this particular polyhedron ξ_0 and let $\Delta\xi = \xi - \xi_0$ be the position vector of ξ relative to ξ_0 in this lattice It will be shown that the term in (15) coming from $\xi_0 = (2\pi/d)(l_0, m_0, n_0)$ alone gives a contribution to J_2 which is not negligibly small

In the first place notice that the contributions of the more distant points fall off quite rapidly on account of the second factor in (15), and that as ξ nears ξ_0, J_2 rises rapidly on account of the contribution of ξ_0 Turning attention to the trigonometric functions, it must be remembered that ρ_0 does not have an accurate value. It is to a certain extent arbitrary both in the integration of (III), when it was first introduced, and in the cut-off introduced in obtaining $F(x, y, z)$. The essential effect must remain, when ρ_0 goes through the limited range of arbitrary values. Now when (l, m, n) is so far away from ξ so that $\sqrt{[(2\pi l/d - \xi_x)^2 + (2\pi m/d - \xi_y)^2 + (2\pi n/d - \xi_z)^2]}$ is large compared with $\pi/(2\rho_0)$, a small fractional change in ρ_0 would completely reverse the sign of this particular term Contributions from such terms are obviously illusory and to be discarded For present purposes, it may be pointed out that it can easily be verified that a change of $\frac{1}{3}\rho_0$ would completely reverse even the signs of the contributions of the nearest neighbours of ξ_0, so long as attention is confined to points in an inner region of the polyhedron of one-third its linear dimension (One takes $2\rho_0 \simeq d$, which is the proper value to take for ρ_0, as will be shown later) It will be seen later that this inner region comprises all that will interest us Therefore the term due to ξ_0 in (15) need only be considered, if an uncertainty of $\frac{1}{3}\rho_0$ in ρ_0 be accepted. With this simplification one obtains in place of (15)

$$J_2 = NSf^2(p-p^2)\frac{8\pi^2c^2}{d^3\rho}|\xi|^2\frac{1}{(\Delta\xi_x)^2+(\Delta\xi_y)^2+(\Delta\xi_z)^2}\left\{\cos 2\rho_0\sqrt{[(\Delta\xi_x)^2+(\Delta\xi_y)^2+(\Delta\xi_z)^2]}\right.$$
$$\times\left(\frac{5}{3}-\frac{[\xi_x\Delta\xi_x+\xi_y\Delta\xi_y+\xi_z\Delta\xi_z]^2}{[\xi_x^2+\xi_y^2+\xi_z^2]\,[(\Delta\xi_x)^2+(\Delta\xi_y)^2+(\Delta\xi_z)^2]}\right)+\frac{\sin 2\rho_0\sqrt{[(\Delta\xi_x)^2+(\Delta\xi_y)^2+(\Delta\xi_z)^2]}}{2\rho_0\sqrt{[(\Delta\xi_x)^2+(\Delta\xi_y)^2+(\Delta\xi_z)^2]}}$$
$$\left.\times\left(5\frac{[\xi_x\Delta\xi_x+\xi_y\Delta\xi_y+\xi_z\Delta\xi_z]^2}{[\xi_x^2+\xi_y^2+\xi_z^2]\,[(\Delta\xi_x)^2+(\Delta\xi_y)^2+(\Delta\xi_z)^2]}-\frac{5}{3}\right)\right\}, \qquad (16)$$

which is in a form suitable for interpretation in a similar manner to the first part of (12), which has been called J_1. These terms and the magnitudes of the effects will now be discussed

Consider first J_1:

$$J_1 = Sf^2\sum_n\sum_m\left[1-(1-\delta_{nm})(p-p^2)\frac{4\pi c^2}{3\rho_0 v}|\xi|^2\right]e^{i\xi\cdot(\mathbf{a}_n-\mathbf{a}_m)(1+4\pi cp/(3v))}$$

Before this term can be interpreted, the Laue effect must be taken into consideration. To do this, it must be remembered that in deducing (7), the terms of (4) were divided into groups of terms with the same $\mathbf{a}_n - \mathbf{a}_m$ and their mean values substituted for the terms in the same group. Now it is necessary to subdivide each group into three smaller groups, namely, $\mathbf{a}_n, \mathbf{a}_m$, both occupied by foreign atoms, one occupied by a foreign atom, and both occupied by solvent atoms. The ratio of the numbers in

these smaller groups is $p^2:2p(1-p):(1-p)^2$ Since they differ only in the scattering factor, it follows that this consideration leads only to a modification of f^2 in J_1, i e.

$$J_1 = S[(1-p)f_1+pf_2]^2 \sum_n \sum_m \left[1-(1-\delta_{nm})(p-p^2)\frac{4\pi c^2|\xi|^2}{3\rho_0 v}\right] e^{i\xi.(a_n-a_m)(1+4\pi cp/(3v))}$$

The constant c can be eliminated For, according to the approximate experimental rule of Vegard, the lattice constant of an alloy changes with concentration p linearly in the following manner

$$d = d_1+p(d_2-d_1) = d_1+p\Delta d, \tag{17}$$

where d_2, d_1 are the lattice constants of the pure specimens of the solute and solvent metals respectively. But it follows from analysis that the lattice constant as revealed by X-ray diffraction should differ from that of the pure solvent metal by the factor $1+(4\pi cp)/(3v)$. Hence, by a simple comparison,

$$\frac{4\pi c}{3v} = \frac{\Delta d}{d_1}, \tag{18}$$

a relation determining c.

Considering a cubic face-centred lattice, ρ_0 can be given a precise meaning. It was originally introduced in evaluating (II) by the crude method given in equation (10) As has been mentioned, (II) has been accurately evaluated in simple cases, and the result is known in particular for a cubic face-centred lattice In fact, making use of a table given by Born & Misra (1940) it is easy to find that the accurate value for (II) is

$$\frac{2\,11\,c^2|\xi|^2}{(d/2)^4} = \frac{33\,7\,c^2|\xi|^2}{d^4}.$$

Remembering that $v = \frac{1}{4}d^3$, one finds by comparison with (10) almost exactly

$$2\rho_0 = d. \tag{19}$$

This will serve as a convenient value to be used for ρ_0 in (III) as given by equation (11). In fact the only reason why the crude method has been applied to expression (II) is to obtain this relation, which will be used for the discussion of J_2

After c and ρ_0 have been eliminated by means of (18) and (19), it is found that J_1 leads to a weakening of the interference maxima by the factor

$$\frac{1}{f_1^2}[(1-p)f_1+pf_2]^2\left[1-6\pi p(1-p)\left(\frac{\Delta d\sin\theta}{\lambda}\right)^2\right]$$

$$\simeq \frac{1}{f_1^2}[(1-p)f_1+pf_2]^2 e^{-6\pi p(1-p)[(\Delta d/\lambda)\sin\theta]^2}$$

The linear expression has been replaced by the exponential function for convenience in discussing temperature effects This is permissible, within the accuracy of our approximation. (Incidentally, the linear and the exponential expressions probably give the limits to the actual value)

The experimental measurement of the intensity of interference maxima is, however, usually made in terms of the 'integrated reflexion' $W\omega/P$ (Compton 1926), which is proportional to the square of the atomic density So when considering this quantity, the modification factor then becomes

$$\left(1-6p\frac{\Delta d}{d_1}\right)\frac{1}{f_1^2}[(1-p)f_1+pf_2]^2\,e^{-6\pi p(1-p)[(\Delta d/\lambda)\sin\theta]^2}, \tag{20}$$

where (17) has again been employed to estimate the change in atomic density

In (20), the first factor is a constant, and the second factor can be calculated for any particular reflexion from fairly reliable theoretical values of f_1 and f_2 (Randall 1934, Barrett 1943) So the last factor can be tested if its effect is not too small Table 1 shows the values of the last factor for different orders of reflexion from (1, 0, 0) and (1, 1, 1) planes of Cu-Au solutions Four cases of the solutions are included, namely, 10 and 20 % solutions of Au dissolved in Cu and vice versa

TABLE 1

order of reflexion		(200)	(400)	(600)	(800)	(111)	(222)	(333)	(444)
Au in Cu	10 %	0 97	0 89	0 77	0 63	0-98	0 92	0-82	0 71
	20 %	0 95	0 81	0 63	—	0 96	0 86	0 71	—
Cu in Au	10 %	0 98	0 91	0 81	0 70	0 98	0 94	0 86	0 76
	20 %	0 96	0 85	0 70	—	0 97	0 89	0 77	—

So far the temperature effect has been entirely left out of consideration Simple considerations serve to show that when temperature effect is also taken into account it is merely necessary to modify the second factor in the previous expression into

$$\frac{\left|(1-p)f_1\exp\left[-\frac{8\pi^2\overline{(u_1')^2}\sin^2\theta}{\lambda}\right]+pf_2\exp\left[-\frac{8\pi^2\overline{(u_2')^2}\sin^2\theta}{\lambda^2}\right]\right|^2}{\left|f_1\exp\left[-\frac{8\pi^2\overline{(u_1)^2}\sin^2\theta}{\lambda^2}\right]\right|^2},$$

where $\overline{(u_1)^2}$, $\overline{(u_1')^2}$ and $\overline{(u_2')^2}$ are the mean values of the squares of the thermal displacement components perpendicular to the reflexion planes, for the solvent atoms in the pure metal, the solvent atoms in the solution and the solute atoms in the solution. It is obviously a difficult task to estimate the two latter quantities with accuracy, for it involves an estimation of the difference in the vibrations of the different atoms for various modes of vibrations of the lattice. It is possible, however, to gain an idea about the magnitude of the effect by examining the two extreme modes of vibration:

(1) *Long elastic waves.* In this case one would expect $\overline{(u_2')^2}\simeq\overline{(u_1')^2}$ The exponential factor can be taken out of the expression So the temperature effect only introduces the additional factor

$$\exp\left\{\frac{16\pi^2\sin^2\theta}{\lambda^2}[\overline{(u_1)^2}-\overline{(u_2')^2}]\right\},$$

which will combine with the factor describing the distortion effect. An estimate of this effect may be formed by contrasting the values

$$\left.\begin{aligned} 16\pi^2\overline{u^2} &= 0{\cdot}91\times10^{-16} \text{ in Au} \\ 16\pi^2\overline{u^2} &= 1\cdot10\times10^{-16} \text{ in Cu} \end{aligned}\right\} T = 300^\circ\text{K},$$

$$\left.\begin{aligned} 16\pi^2\overline{u^2} &= 0\cdot32\times10^{-16} \text{ in Au} \\ 16\pi^2\overline{u^2} &= 0{\cdot}42\times10^{-16} \text{ in Cu} \end{aligned}\right\} T = 89^\circ\text{K},$$

and guess that $16\pi^2[\overline{(u_1')^2}-\overline{(u_2')^2}]$ is of the order $\lesssim 0\cdot05\times10^{-16}$ for $T = 300^\circ$K and $\lesssim 0{\cdot}02\times10^{-16}$ for $T = 89^\circ$K. The corresponding quantity in the distortion effect is $0\cdot4\times10^{-16}$ for $p = 10\,\%$ and $0{\cdot}8\times10^{-16}$ for $p = 20\,\%$.

(ii) *High-frequency vibrations* In this case neighbouring atoms vibrate almost independently. To form a very crude idea of the magnitude of the effect, consider each kind of atom to vibrate in the same way as in its own metallic lattice. Since the effect increases with the order of reflexion, the second factor is calculated both with and without the temperature effect for the reflexion (800). The results are given only for Au dissolved in Cu, for in the reverse case the difference is completely negligible.

TABLE 2

	10 %		20 %	
	89°	300°	89°	300°
with temperature effect	1·68	1·84	2·55	2·70
without temperature effect	1·62	1·62	2·40	2·40

The difference amounts to $\sim 10\,\%$ at the most.

Now proceed to discuss J_2 as given by (16). Consider in the first place the points at which $\xi_x\Delta\xi_x+\xi_y\Delta\xi_y+\xi_z\Delta\xi_z$ vanishes. In the inner region, these are points lying almost in the plane perpendicular to ξ_0 and passing through the centre of the polyhedron. For these points, when $2\rho_0\sqrt{[(\Delta\xi_x)^2+(\Delta\xi_y)^2+(\Delta\xi_z)^2]}\lesssim 0{\cdot}3$, it is easy to see from (16) that the intensity practically vanishes. On the other hand, proceeding along the direction of ξ_0 from the point ξ_0, the intensity will be strongest. In what follows, the intensity at various points along this line will be considered, which will allow a comparison of the magnitude of the effect with the corresponding thermal effect. Then if a quantity τ be defined such that $2\pi\tau = \sqrt{[(\Delta\xi_x)^2+(\Delta\xi_y)^2+(\Delta\xi_z)^2]}$ on this line and c and ρ_0 eliminated from (16) by means of (18) and (19), then

$$J_{2\parallel} = 3NSf^2(p-p^2)\left(\frac{\Delta d}{d}\right)^2\left(\frac{\sin^2\theta}{\lambda^2}\right)\frac{1}{\tau^2}\left\{\cos 2\pi d\tau + 5\frac{\sin 2\pi d\tau}{2\pi d\tau}\right\} \tag{21}$$

Since for the present purpose one is only interested in roughly comparing the relative orders of magnitude of the effects, it is found convenient to use the result of an earlier paper by Zachariasen (1940) for the intensity of thermal diffuse maxima. In that paper, he has used the same velocity for elastic waves of all directions and

polarizations His results will be used with a slight modification. Two velocities, v_t for transverse waves, v_l for longitudinal waves, are employed When the linear dimension of the crystal is $\lesssim 10^{-4}$ cm., then for all relevant values of τ, the slightly modified form of Zachariasen's result is

$$J_{\mathbf{s}}' = NSf^2 \frac{64h}{3m_a}\left(\frac{\sin^2\theta}{\lambda^2}\right)\left\{\frac{1}{v_l\tau}\left(\frac{1}{\exp\left[\frac{hv_l\tau}{kT}-1\right]}+\frac{1}{2}\right)+\frac{2}{v_t\tau}\left(\frac{1}{\exp\left[\frac{hv_t\tau}{kT}-1\right]}+\frac{1}{2}\right)\right\}, \quad (22)$$

which corresponds to (21) in the present case, where m_a is atomic mass.

The numerical values for $J_{\mathbf{s}\parallel}$ and $J_{\mathbf{s}}'$ will be given for some relevant values of τ for the case of copper as solvent. If one allows for a diffusion of the primary maxima to the extent of $\sim 0{\cdot}5°$, it is easy to find that only values of τ greater than about $1/100\lambda$ are of interest. Assuming a value of $\lambda \sim 1\ 5$ A, $J_{\mathbf{s}\parallel}$, $J_{\mathbf{s}}'$ may therefore be calculated starting with $\tau = 1/40d$ In table 3, the factor $NSf^2(\sin^2\theta/\lambda^2)$ which appears in both (21) and (22) has been omitted $J_{\mathbf{s}}'$ has been given for two temperatures $T = 300$ and 100° K.

TABLE 3

τ		$1/40d$	$1/30d$	$1/20d$	$1/10d$	$1/5d$
$\frac{J_{\mathbf{s}\parallel}\times 10^{13}}{p(1-p)}$		5 8	3 3	1·4	0 34	0·06
$J_{\mathbf{s}}'\times 10^{13}$	300° K	6 5	3·8	1 5	0 39	0·10
	100° K	2 1	1·2	0 5	0·13	0 03

In the calculation of $J_{\mathbf{s}}'$, the velocities of the elastic waves are obtained from the isotropic elastic constants obtained by averaging the actual elastic constants over all directions. In confirmation of some remarks made previously, it is observed that at $\tau = 1/5d$, the intensity is reduced to only about a hundredth that at $\tau = 1/40d$ and $1/5d$ is still within the inner region of the polyhedron where (21) should hold.

As far as the values in table 3 indicate, the effect can be comparable with the thermal effect only at very low temperatures ($< 100°$ K) if the solution is still to be dilute But the thermal effect is obviously greatly affected by the atomic mass and to an even greater extent by the actual values of the velocities of elastic waves So it cannot be concluded from the values shown that this effect will always be masked by the thermal effect. In fact, the simple result of Zachariasen quoted here completely neglects the dependence of $J_{\mathbf{s}}'$ on the direction of $\Delta\boldsymbol{\xi}$, yet any experimental possibility of detecting the effect probably depends on setting $\mathbf{n}_1$ in a direction exhibiting small thermal effect It is best, however, to abstain from such specific investigations and leave the analysis as it stands

In conclusion, I should like to express my thanks to Professor N. F Mott, F R.S., for suggesting this problem and for his constant guidance.

References

Barrett 1943 *Structure of metals*, p 529. London. McGraw-Hill
Born 1940 *Proc. Camb. Phil. Soc* **36**, 1
Born & Misra 1940 *Proc. Camb Phil Soc* **36**, 466.
Compton 1926 *X-rays and electrons*, p 142 New York· Van Nostram.
Debye 1914 *Ann. Phys., Lpz*, **43**, 49.
Lonsdale 1942 *Proc Phys. Soc.* **54**, 314.
Randall 1934 *The diffraction of X-rays and electrons by amorphous solids, liquids and gases*, p. 263. London Chapman & Hall.
Von Laue 1918 *Ann. Phys*, *Lpz*, **56**, 497.
Zachariasen 1940 *Phys Rev* **57**, 597.

The adsorption of vapours on mercury
III. Polar substances

By C. Kemball, *Department of Colloid Science, The University, Cambridge*

(*Communicated by E K Rideal, F R S.—Received* 25 *October* 1946)

Water, acetone and the normal alcohols from methyl to hexyl have been adsorbed on mercury All substances gave reversible adsorption and, with the exception of water, gaseous films were formed at low pressures. Methyl and ethyl alcohols showed the adsorption of a second layer at higher pressures, the double layer having half the co-area of the original monolayer Actone gave rise to a double layer and finally a triple layer (with one-third of the original co-area) The property of multilayer formation was thought to bo one of small partially polar molecules For the gaseous films of *n*-butyl, *n*-amyl and *n*-hexyl alcohols the co-areas and the thermodynamic data indicated that the molecules were lying flat on the surface These three substances showed two-dimensional condonsation to liquid films at higher pressures This phase change was accompanied by an increase of entropy which led to a decrease of the surface-vapour pressure with rise of temperature The large entropy and heat of adsorption of water were taken as evidence for the association of the adsorbed water molecules and this probably occurred, to some extent, with methyl alcohol as well The heat of adsorption of acetone was smaller than expected for a substance with a large dipole moment

Introduction

In Parts I and II (Kemball & Rideal 1946; Kemball 1946*a*) details were given of the adsorption of some non-polar substances on mercury. The thermodynamic quantities for the formation of the monolayer were determined in each case, and the monolayer was found to be stable up to high values of p/p_0, p being the vapour pressure and p_0 the saturation vapour pressure. The evidence for the formation of a second adsorbed layer on top of the first was not conclusive, nor was a clear case of two-dimensional

condensation discovered. The work described in the present paper was undertaken with the following problems in mind:

(1) The determination of the heats of adsorption of some polar molecules for comparison with the results already obtained with benzene, toluene and *n*-heptane.

(2) To confirm or disprove the fact that water vapour was not adsorbed on mercury which was reported by Cassel & Salditt (1931)

(3) The search for substances which would form second and third layers at high values of p/p_0. There was considerable hope of finding such substances among polar compounds because some behaviour of this type had been shown with toluene.

(4) (*a*) To discover the effect of increasing the chain length on the heat of adsorption and on the orientation of the molecules at the surface

(*b*) To increase the chain length with a view to obtaining two-dimensional condensation

A brief outline of the published work on the adsorption of vapours on mercury was given in Part I (Kemball & Rideal 1946)

Experimental

The apparatus and the measurement of the surface tension of mercury in the presence of vapours has been described elsewhere (Kemball 1946*b*). The purification and the data for the determination of the vapour pressure from temperature, of the substances adsorbed, were as follows:

Methyl alcohol. Commercial methyl alcohol was distilled three times; b p 64·2° C/760 mm. Vapour pressures were calculated from the $\log p$ against $1/T$ relation obtained from the data of Ramsay & Young (1887), Flock, Gunnings & Holton (1931) and Mundel (1913*a*)

Ethyl alcohol After distillation the alcohol was dried over potassium carbonate, $\eta_D^{20} = 1{\cdot}3625$. The results of Mundel (1913*a*) and of Ramsay & Young (1886) were used to calculate the vapour pressure

n-Propyl alcohol. The alcohol was dried over lime and distilled three times, b p. 96·5° C/760 mm., $\eta_D^{20} = 1{\cdot}3860$. A $\log p$ against $1/T$ relation was obtained from the work of Mundel (1913*b*) and Schmidt (1891).

n-Butyl alcohol The alcohol was distilled, b p. 116 6° C/765·8 mm., $\eta_D^{20} = 1{\cdot}3995$. The best straight line was fitted to the $\log p$ against $1/T$ data from 30 0 to 60·2° C, given by Gilliland & Sherwood (1934). For lower temperatures extrapolation of this line was used

n-Amyl alcohol I am grateful to Professor N K Adam, F R S., for the specimens of amyl and hexyl alcohol used in the present research. The specimens had been prepared by the Eastman Kodak Company After distillation the value of $\eta_D^{20} = 1{\cdot}4099$. The information available for the determination of the vapour pressure was most unsatisfactory. A $\log p$ against $1/T$ relation was obtained from the results of Grassi (1888), who did not specify which isomer of the alcohol was used. By comparison with the corresponding lines for *n*-butyl and *n*-hexyl alcohols the

gradient of this line was found to be correct but the whole line appeared to be displaced by as much as 0·17 unit on the logarithm to base ten scale

n-Hexyl alcohol. The Eastman Kodak specimen was distilled and the central fraction used, $\eta_D^{20} = 1{\cdot}4170$. The vapour pressure determinations of Hovorka, Lankelma & Stanford (1938) were employed

Water Identical results were obtained with distilled and tap water. The vapour pressures were taken from the *Handbook of Physics and Chemistry*, 1943–4.

Acetone. The acetone was dried over calcium sulphate and distilled, b p 56·5° C/760 mm, $\eta_D^{20} = 1{\cdot}3591$. The values of the vapour pressure given by Drucker & Kangro (1915) and Sameshima (1918) were found to be in agreement and were used to determine the vapour pressure from temperature.

Experiments were conducted with the mercury at two temperatures, 25 and 50° C With the higher alcohols it was necessary to wait for as long as 2 hr at the lower pressures before equilibrium was attained throughout the apparatus This length of time was required because the alcohol vapour was diffusing through a pressure

TABLE 1 METHYL ALCOHOL

temp °K of alcohol	p (mm.)	F (dynes/cm)	temp °K of alcohol	p (mm)	F (dynes/cm)
	25° C, surface tension of mercury 484 2 dynes/cm.				
208 7	0 144	5 0	255 1	8 55	41 8
214 4	0 260	7 0	255·7	8 89	42 8
218 5	0 394	9 7	258 6	10 94	46 0
222 6	0·583	12·4	264 6	16 37	52·4
227·5	0 918	14 0	266 0	18 03	54 9
220 2	1 16	16 0	270 1	23 39	57 5
232·2	1 40	18·0	271 6	25 70	60 6
235 0	1·77	19·3	272 7	27 67	62 5
237 3	2 14	21 3	275 3	32 21	65 0
239 5	2 58	22 6	279 8	42 36	67 0
242·1	3 18	26 2	283 6	53 21	72·5
245 2	4 06	29 8	289 4	73 62	78 6
249·7	5 72	35·7	291·4	81 65	80 9
	50° C; surface tension of mercury 479 0 dynes/cm				
230 2	1·17	5 0	272 7	27·54	27 7
234 7	1·73	7 0	276 8	35 40	31 2
239·5	2·58	8 7	277 4	36 81	32 2
243·3	3 47	11 0	277 7	37 41	33 2
246 2	4 39	12 0	283 1	51 88	38 0
249 4	5·61	13 9	283 4	52 48	38 6
252 1	6·86	15 2	286 5	62·37	40 9
254 6	8 24	17 2	286 7	62 95	41 8
256·7	9 53	18 5	288 4	69·18	43 2
259 3	11 46	20 2	288 8	70 79	43·4
261·5	13 34	21 1	293 4	91 20	47 9
264 7	16 56	22·5	294·3	96 16	48 8
267 4	19·82	24 7	297 2	111 2	51 7
269·8	22 06	26 4			

of mercury vapour of similar magnitude All the isotherms were completely reversible. The experimental results for the alcohols from methyl to *n*-amyl, and for water are given in tables 1 to 6 Those for acetone and *n*-hexyl alcohol are given graphically (figures 1 and 2) The curves of surface pressure against vapour pressure for methyl

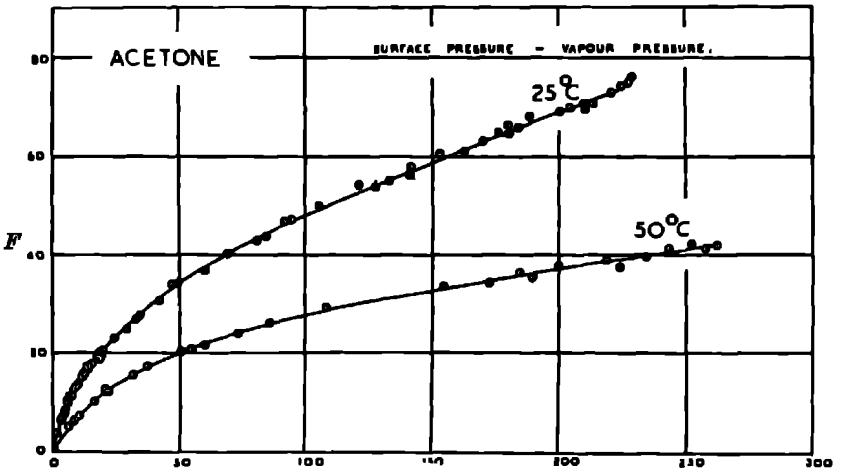


FIGURE 1

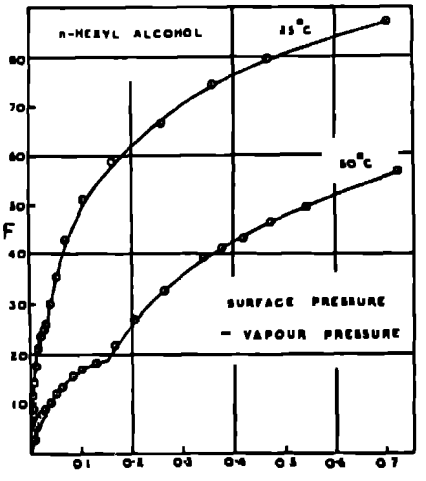


FIGURE 2

and ethyl alcohols were similar to that for acetone (figure 1), and those for *n*-butyl and *n*-amyl alcohols, to the one for *n*-hexyl alcohol (figure 2) Water and propyl alcohol gave results similar to the 50° C curve of acetone, i e. concave to the pressure axis but without the tendency to increase in slope shown by the 25° C curve.

TABLE 2. ETHYL ALCOHOL

temp. ° K of alcohol	p (mm.)	F (dynes/cm)	temp ° K of alcohol	p (mm.)	F (dynes/cm)
	25° C, surface tension of mercury 484·2 dynes/cm.				
226 8	0·213	9 4	258 9	4 17	36 6
231·5	0·368	12 4	263 0	5 69	40 8
233 5	0 447	13 7	266 8	7·57	46 0
235 6	0 550	16·0	269·0	8 93	50 5
239·1	0 762	18 3	271 9	11 02	54 0
244 0	1·19	22 0	274 4	13 12	58 1
247 2	1·58	25·2	276 5	15 28	61·3
249·9	2 00	27 9	279 0	18 07	65 4
251 9	2·36	29 8	283 0	23 66	71 0
254·5	2·93	32 4	285 7	28·38	75·3
257 2	3 64	35 3	290 1	37 15	82 1
	50° C, surface tension of mercury 470 0 dynes/cm				
242 4	1 03	9 6	276 6	15 38	34 8
245·1	1 31	11 3	280 4	19 91	38 9
248 1	1 72	13 3	282 4	22 75	40 5
251·1	2·21	14 9	284·7	28 36	42 8
253 7	2·74	16 9	286 9	30 48	45 0
256 2	3 35	18 9	287·5	31 84	45 0
259 9	4 53	21 5	290·6	38·55	47 5
266 9	7 66	26 7	293·2	44·08	51 4
269 5	9·29	29 3	294·9	49 89	52 6
271 8	10 96	31 5	297 0	56 49	54·6
273 5	12 36	33 2	298 7	62 81	57 4

TABLE 3. *n*-PROPYL ALCOHOL

temp ° K of alcohol	p (mm)	F (dynes/cm)	temp ° K of alcohol	p (mm.)	F (dynes/cm.)
	25° C, surface tension of mercury 484·2 dynes/cm.				
229 1	0 0706	7·0	268 5	2 59	43 4
233 1	0 107	9·0	271 3	3·21	46 0
238·4	0·181	13 4	278·7	5 59	54·6
241 8	0·251	15·4	281 3	6 67	57 8
246 8	0 406	19·3	283·1	7 64	60·3
250·6	0 573	23 6	286·7	9 75	65·0
257 3	1 03	32 1	289 2	11·69	67·9
262 5	1 60	36·0	291·6	13·55	71 3
265·4	2·04	39 5			
	50° C; surface tension of mercury 479 0 dynes/cm				
227·4	0 0589	1 7	267 6	2·41	22 8
233·2	0 109	3·3	271 0	3·14	25·4
238·4	0 181	5 0	273 6	3·83	27·7
250·0	0 545	11·0	277·5	5 12	31·2
252 6	0·686	12·0	280·2	6 22	33 8
256 7	0·900	13 9	282 3	7 19	35·4
259 7	1 27	15 9	284 2	8 61	37·7
262·1	1·55	17 9	287·5	10·12	40·5
264·8	1·94	20·8	289·3	11·72	42·1

TABLE 4. *n*-BUTYL ALCOHOL

temp. ° K of alcohol	p (mm.)	F (dynes/cm)	temp ° K of alcohol	p (mm)	F (dynes/cm)
	25° C, surface tension of mercury 484·2 dynes/cm.				
214 6	0 00290	3·4	263 3	0 504	40·8
219 7	0 00553	5 0	264·6	0 563	42·1
227 3	0 0138	9 7	267 1	0 697	45 3
229·2	0 0172	10 7	267 2	0 703	46 0
231 5	0 0222	12 0	269·1	0·823	47 9
235 8	0 0356	15 7	271 0	0 961	51 7
239 3	0 0517	18·7	273·6	1·19	55 9
245 2	0 0941	23 3	277 0	1·55	59 4
248 2	0·127	27·2	278·9	1·82	63·5
253 3	0 206	31 8	279 7	1·91	65 0
257·2	0·294	35 0	283 2	2 49	68·8
258 0	0 316	36 0	285 0	2 85	71·3
260 4	0 391	37 9	286 5	3 18	77 2
	50° C; surface tension of mercury 470 0 dynes/cm				
238 1	0·0455	4 3	271 7	1·01	24·4
240 1	0 0561	5 3	276·8	1 53	29·3
244 3	0 0861	6 7	279 2	1·84	32 5
248 8	0·135	8 3	278·9	1 80	32 8
252 2	0 186	10·6	282 1	2 29	36·4
254 8	0 236	12 6	283 9	2 62	39 3
257 1	0·291	14 6	285 4	2 93	41 8
258 8	0·339	15 2	287 3	3 37	44 4
260 2	0·384	15 9	288 8	3 75	46 3
261 7	0·439	17 2	290 3	4 17	47 9
265·2	0 503	20 2	291 4	4 51	50 1
266 2	0 642	20 5	291 9	4 67	50·4
268 3	0 770	21 5	293 7	5 29	52 0
268 9	0 810	21 8	296 2	6 27	54 9

INTERPRETATION OF THE EXPERIMENTAL RESULTS

The methods of interpretation which had proved successful with the non-polar substances were used. The first of these was the fitting of the Volmer equation (1) to the experimental results, on the assumption that one was dealing with a mobile film. The plot of $\log p/F$ against F was constructed in each case to test the relation (2) derived from the Volmer equation and the Gibbs equation

The only attempt that was made to apply the Langmuir equation was in the case of water where there was reason to believe that the molecules were immobile on the surface This will be discussed in detail below

In order to understand the phase changes occurring in nearly all the adsorption isotherms at higher pressures, the $F - \log p$ curves were drawn. The gradient of this type of curve

$$F(A-b) = kT, \tag{1}$$

$$\ln \frac{p}{F} = \frac{bF}{kT} + c, \tag{2}$$

TABLE 5 *n*-AMYL ALCOHOL

temp °K of alcohol	p (mm)	F (dynes/cm)	temp °K of alcohol	p (mm)	F (dynes/cm.)
	25° C, surface tension of mercury 484 2 dynes/cm				
226 9	0 00642	11 0	260 3	0 200	40 2
231·1	0 0103	13 7	265 3	0 311	47 0
235 0	0 0160	16 0	268 3	0 404	51 7
237 8	0 0217	18 3	272 7	0 557	58 7
240 7	0 0296	20 6	277 0	0 826	65 4
243 6	0·0401	23 3	282 6	1·28	72·2
247 7	0·0608	26 0	284 4	1·47	76 0
249·7	0·0741	28 5	287 4	1 84	70 4
253·8	0 116	31 4	289·2	2 09	81 2
257·8	0 160	37 0	292 0	2 56	84 9
	50° C; surface tension of mercury 479 0 dynes/cm				
235 0	0·0160	4 3	270 2	0 474	23 1
241 0	0 0306	6 0	271 5	0 527	24 1
246 3	0 0527	8 3	275 4	0 726	30 6
248 5	0 0658	10 0	278 8	0 953	35 4
251·3	0 0863	11 3	279 5	1·01	37 0
253 9	0 111	13 0	282 7	1 29	42·1
256 4	0 140	14 6	285 2	1 56	47·2
259·9	0 104	16 9	289 0	2 07	51 7
262·2	0·238	18 2	201·6	2 50	55·5
264 5	0 290	19 8	294 4	3 04	60·5
266 2	0·337	20 8			

TABLE 6 WATER

temp °C of water	p (mm.)	F (dynes/cm)	temp °C of water	p (mm)	F (dynes/cm)
	25° C; surface tension of mercury 484·2 dynes/cm.				
−15 7	1 16	8 0	9 8	8 88	31·1
−14 0	1·36	8 4	10·6	9 59	32·7
−13 3	1 45	9 4	11 3	10·04	32·7
−12·8	1 52	9 4	13 3	11 45	34·7
−11 0	1 66	10·0	13 1	11 32	35 0
− 8 3	2 27	13 0	14 2	12 14	36·0
− 6 2	2 72	16 0	17 1	14 62	38 3
− 3 4	3 51	18 7	17 8	15 09	39 2
− 2 1	3 85	19 3	19 9	17 43	42 5
0 0	4·58	21 6	20 6	18 20	43 1
3 4	5 85	24 3	22·7	20 69	44 7
6 3	7·16	26 9	23 4	21 60	45 7
9 1	8·67	29 8	23 3	21 45	46 0
	50° C; surface tension of mercury 479·0 dynes/cm.				
5 7	6 87	5 3	21 4	19 37	13 3
9 5	8 90	6 3	23 6	21 85	13 9
11·3	10·04	7·3	25 8	24 62	14·9
14 5	12 38	9 0	28 4	29 02	16·2
17 2	14·72	10·3			

gives the amount on the surface according to the Gibbs equation In this way it was possible to obtain force-area and (force) (area)-force curves for the polar substances on mercury.

The primary gaseous film

In all cases the first region of the adsorption isotherm gave a straight line within the limits of experimental error when plotted as $\log p/F$ against F, indicating a gaseous film obeying a Volmer equation Accordingly, the free energies, heats and entropies of adsorption were calculated for each of the substances by the method outlined in Part I (Kemball & Rideal 1946) The values obtained are shown in table 7. The errors given are exclusive of errors in the vapour pressure values In the case of *n*-amyl alcohol the additional error would probably amount to 250 cal. in ΔG, less than 1 entropy unit in ΔS and negligible error in ΔH and b, the co-area. Some examples of the $\log p/F$ against F curves obtained are shown in figures 3–7 Ethyl alcohol gave results similar in form to those of methyl alcohol (figure 3), where the second region of each isotherm was also a straight line on this plot Some deviation from the linear relation can be observed near the high pressure end of the propyl alcohol isotherm at 25° C (figure 4). *n*-Butyl and *n*-hexyl alcohols yielded curves similar to those shown in figure 5 for *n*-amyl alcohol. The second regions, unlike those for the first two alcohols, were no longer straight lines The transition pressures became lower the longer the hydrocarbon chain in the molecule The results for water (figure 6) require no comment and those for acetone (figure 7) will be discussed below under the section on the formation of second layers.

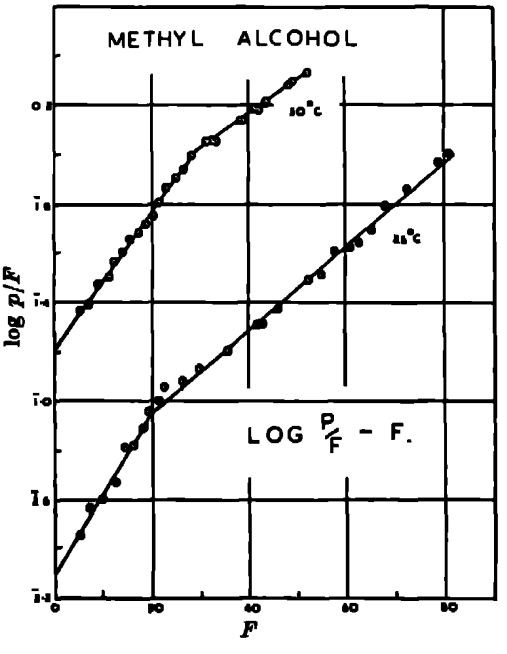


FIGURE 3

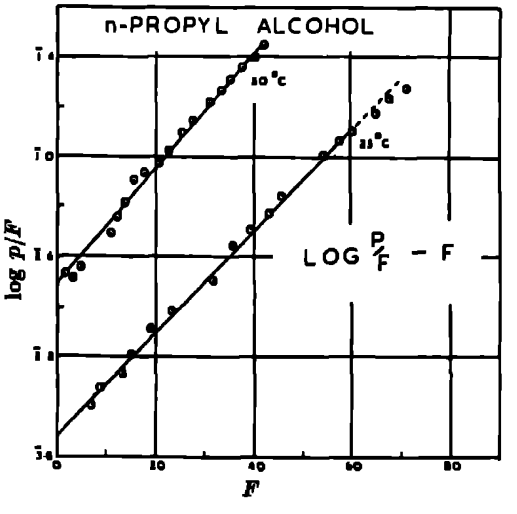


FIGURE 4

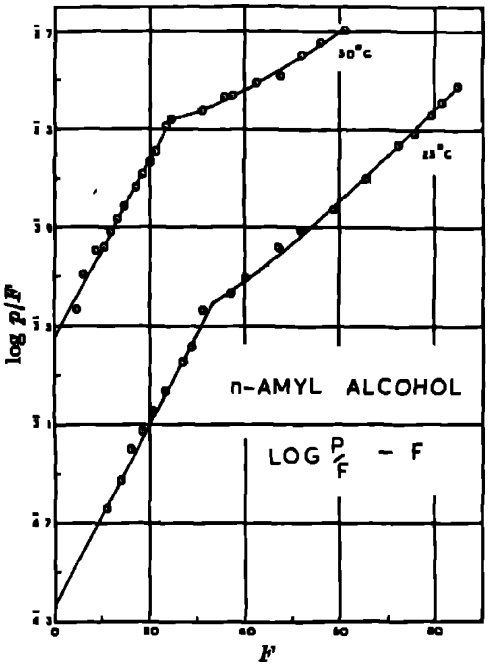


FIGURE 5

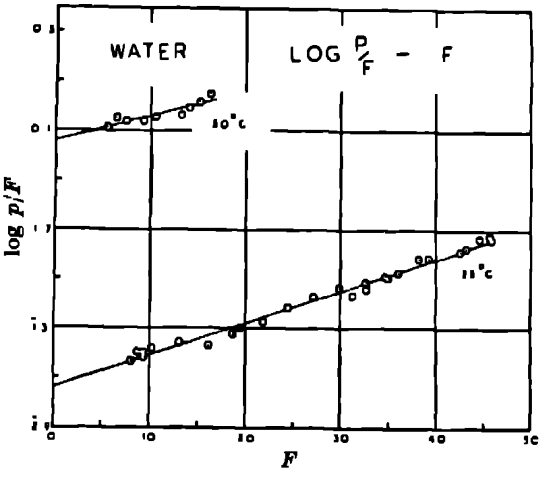


FIGURE 6

TABLE 7. DETAILS OF THE PRIMARY ADSORBED LAYER

substance	temp °C	co-area	ΔG cal /g mol	ΔH kcal /g mol	ΔS cal /deg.mol
methyl alcohol	25	31·3 ± 1·0	− 7879 ± 22	−15·9 ± 0·43	−27·2 ± 1·3
	50	28·4 ± 0·4	− 7199 ± 10		
ethyl alcohol	25	22·9 ± 0·3	− 8057 ± 11	−11·4 ± 0·26	−11·2 ± 0·8
	50	23·4 ± 0·2	− 7778 ± 8		
n-propyl alcohol	25	19·1 ± 0·2	− 8445 ± 12	−10·8 ± 0·35	− 7·9 ± 1·1
	50	23·1 ± 0·4	− 8248 ± 16		
n-butyl alcohol	25	28·1 ± 0·3	− 9817 ± 10	−18·1 ± 0·34	−27·8 ± 1·1
	50	29·8 ± 0·8	− 9122 ± 17		
n-amyl alcohol	25	35·2 ± 0·7	−10506 ± 23	−19·2 ± 0·52	−29·0 ± 1·6
	50	37·0 ± 0·7	− 9780 ± 16		
n-hexyl alcohol	25	42·8 ± 0·6	−11231 ± 15	−21·2 ± 0·59	−33·4 ± 1·8
	50	45·0 ± 0·7	−10396 ± 20		
water	25	12·3 ± 0·2	− 6847 ± 8	−17·6 ± 0·36	−35·9 ± 1·1
	50	10·3 ± 1·2	− 5949 ± 19		
acetone	25	18·8 ± 0·3	− 6150 ± 10	− 7·45 ± 0·22	− 4·4 ± 0·7
	50	20·5 ± 0·2	− 6041 ± 8		

The state of adsorbed water

In the case of water the entropy of adsorption was so large that it seemed probable that the molecules were immobile on the surface Presumably any particular water molecule remained in contact with the same mercury atoms for a relatively long period and the complex thus formed was able to move about the surface. It is not clear, therefore, whether the Langmuir equation, which is based on the assumption of adsorption on to fixed sites, would be expected to fit the experimental results more accurately than the Volmer equation. Complete immobility would not be

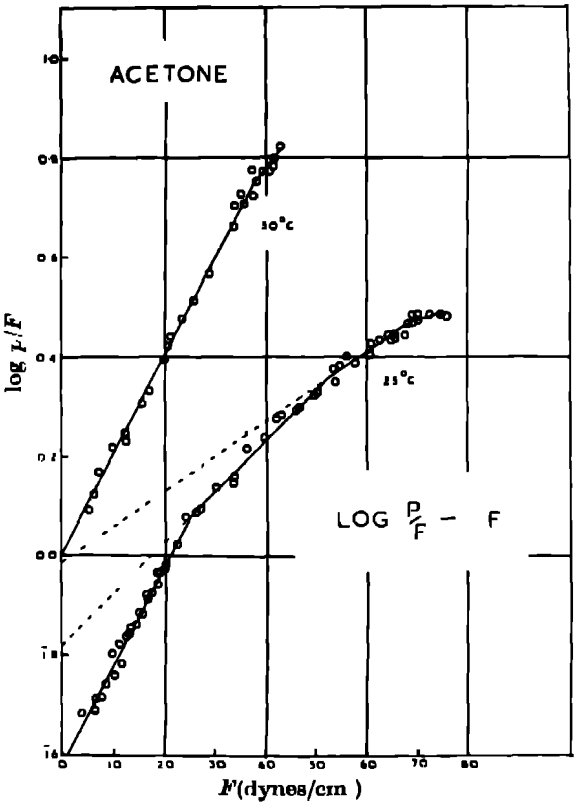


FIGURE 7

expected for adsorption on to a liquid substrate The lower part of the $\log p/F$ against F graph for water (figure 6) did show some tendency to curve, and it was mentioned in Part I that the Langmuir isotherm gave curved results at low pressures on this type of plot. Accordingly a Szyszkowski equation (implying a Langmuir isotherm) was fitted to the results for water at 25° C and, as can be seen from table 8, the value of the quantity R was resonably constant The calculated Szyszkowski equation was

$$F = 0{\cdot}09417 \times 484{\cdot}2 \log\left(\frac{p}{2{\cdot}356} + 1\right) \qquad (3)$$

In order to compare the fit given by this and the relation (4) which had been calculated by the Volmer method outlined above, the error function (5) was evaluated

$$\log p/F = 0\ 01306F - 0\ 9374, \tag{4}$$

$$\Omega = \frac{\Sigma(y_0 - y)^2}{n - m} \tag{5}$$

y_0 is the calculated value of the observed quantity y, n is the number of observations, and m is the number of variables in the equation fitted The experimental values of F were used to obtain calculated values of $\log p$, and these, together with the observed values of $\log p$, were used in the error function (5), which was evaluated for each of the equations in turn. m was 2 in both cases and the values of n were the same. The equation (3) gave a value of $\Omega = 0\ 000329$, compared with 0·000390 from equation (4), indicating that the Langmuir method of interpretation was slightly more accurate than the Volmer. This behaviour contrasts with that found for the adsorption of *n*-heptane (Kemball & Rideal 1946) The Langmuir equation gave the limiting area per molecule as 20 6 A², compared with 12 3 A² obtained as the Volmer co-area

TABLE 8 ADSORPTION OF WATER AT 25° C

γ/γ_0	F (dynes/cm)	$\log p$	p (mm)	$p_n - p_{n-1}$	$R = \frac{p_n - p_{n-1}}{p_{n-1} - p_{n-2}}$	$\log p$ (calc)
1 00	0 00	—	0 000	—	—	—
0·99	4 84	1 82	0 661	0·661	—	1 815
0 98	9·68	0 184	1 528	0 867	1 304	0 172
0 97	14 53	0 408	2 560	1 037	1 195	0·407
0 96	19·37	0 580	3·802	1 307	1 261	0 592
0 95	24 21	0 748	5 598	1 706	1 372	0·752
0 94	29·05	0 895	7 852	2 254	1 253	0 896
0 93	33 89	1 028	10 67	2 818	1 250	1 029
0 92	38 74	1·153	14 22	3·55	1 260	1 156
0 91	43 58	1 278	18 97	4 75	1 333	1 277
					mean $R = 1\ 277$	

As mentioned above, the entropy of adsorption of water was large, $35{\cdot}9 \pm 1{\cdot}1$ entropy units, and calculation on this topic led to interesting conclusions Accurate values of the entropy of water have been available for some time Gordon (1934) gave 45·101 e u. at 298·1° K and 1 atm The discrepancy between this, and the experimental value of 44·28 e u obtained by Giauque & Stout (1936), has been explained on the postulation of random orientation of the hydrogen bonds in ice. The idea was put forward by Pauling (1935) who showed that the difference between experimental and calculated values should be $R \ln 6/4 = 0{\cdot}806$ e u Gordon's value was confirmed by Giauque & Archibald (1937) using the reaction

$$Mg(OH)_2 \rightleftharpoons MgO + H_2O.$$

The contributions of the vibrations to the entropy, and the effect of the vibrations on the moments of inertia are small at room temperature and may be ignored. Using

the moments of inertia given by Gordon (1934) the sum of the translational and rotational entropy is found to be 45·0, compared with the accurate value of 45·101 e u. A similar calculation for 310·6° K gives 45·4 e.u , which, with an entropy of adsorption of 35·0 ± 1·1, means a value for the adsorbed molecules of 9·5 ± 1·1 e.u.

When the molecules are regarded as fixed on the surface there will be no translational entropy for the adsorbed phase, but there will be a considerable entropy corresponding to the number of ways in which the molecules may be distributed between the available sites. Fowler & Guggenheim (1939) gave the number of distinguishable ways of distributing $N_A, N_B, \ldots$ of molecules of types $A, B,$ on N_S sites as

$$g(N_A, N_B, \quad) = \frac{N_S!}{N_A!\,N_B! \quad (N_S - \Sigma N_A)!},$$

which corresponds to an entropy given by

$$S = k \ln \left[\frac{N_S!}{N_A!\,N_B! \ldots (N_S - \Sigma N_A)!} \right].$$

Considering molecules of one type only, letting $N_A = N$, Avogadro's number, and taking the standard state as providing x sites per molecule, i.e. $N_S = Nx$, then by Stirling's approximation

$$S = R[x \ln x - (x-1) \ln (x-1)] \qquad (6)$$

This relation gives the entropy for one mole adsorbed on to fixed sites in terms of one parameter x, which will depend on the size of the molecules and on the distance between the mercury atoms on the surface.

Using equation (6) and taking the area of a water molecule as 12·0 A^2, which gives $x = 583{\cdot}1$ at the concentration of the standard state, the entropy is found to be 14·6 This is much greater than the experimental value of 9 5 ± 1 1 e u There are two possible explanations of this discrepancy First, when one molecule has been adsorbed there may be some hindrance experienced in the adsorption of a molecule to a neighbouring site due to dipole repulsion, this would mean that the area of each site must be increased and a reduction in x would result. However, to reduce the theoretical entropy to 9 5 e.u the value of x would have to be 45 and the area of each site 155 A^2 Secondly, the molecules may be associated on the surface involving a further reduction in their entropy.

To estimate the possible decrease in entropy if all the water molecules were associated into pairs in the surface layer, an expression given by Fowler & Guggenheim (1939) and due to Chang (1939) may be employed. The number of ways of combining N_A double molecules with N_B single molecules is

$$\left[\frac{\{2(z-1)(N_A/z) - N_B\}!}{(2N_A/z)!\,(2N_A + N_B)!} \right]^{\frac{1}{2}z} \frac{(2N_A + N_B)!}{N_B!} \left(\frac{2}{\sigma_{AA'}} \right)^{N_A},$$

where z gives the number of nearest neighbours, and $\sigma_{AA'}$ is a symmetry number which is equal to 2 if the ends of a double molecule are indistinguishable. Replacing N_A by $\frac{1}{2}N$ and considering N_B as representing empty sites, not single molecules, the

expression can be adapted to the present problem The symmetry number is 2, and the co-ordination number for a mercury surface may be taken as 4. The total number of sites is again taken as xN which therefore replaces $2N_A + N_B$. Using Stirling's approximation the entropy per mole of single molecules is given by

$$S = R[2(x-\tfrac{1}{2})\ln(x-\tfrac{1}{2}) - x\ln x - (x-1)\ln(x-1) + \ln 2]$$

Using $x = 583{\cdot}1$, the entropy is found to be 8 7 e u , which is a considerable reduction from 14·6 and in better agreement with the experimental value of 9·5.

The entropy might be further reduced by assuming association into clusters of more than two molecules and by assuming a larger area for each site Factors which have been neglected and which will cause an increase in the calculated entropy, are the vibrations of the molecules in the force field of the mercury and about the link, presumably a hydrogen bond, holding the pairs or clusters of molecules together These considerations of entropy form strong proof of the association of water on the surface of mercury because the large entropy of adsorption can be explained in no other way

Second layer formation

The experimental results for acetone provided evidence for the formation of more than one adsorbed layer The $\log p/F$ against F curve (figure 6) showed certainly two and probably three straight regions The first discontinuity was sharp but the second was ill-defined Straight lines were fitted to the three regions and the values of the constant b obtained were $18{\cdot}8 \pm 0\ 3$, $9{\cdot}6 \pm 0\ 2$, $6\ 5 \pm 0\ 4$ A^2, indicating that the equations

$$F(A - 18{\cdot}8) = kT, \quad F(A - 9\ 6) = kT, \quad F(A - 6\ 5) = kT,$$

were obeyed in succession Now $\frac{1}{2} \times 18{\cdot}8 = 9{\cdot}4$ and $\frac{1}{3} \times 18\ 8 = 6\ 3$, which suggested that the formation of a second and then a third layer was taking place In analogy with the equation

$$F(A - \tfrac{1}{2}b) = kT,$$

which was suggested in Part 1 for two layers on the surface,

$$F(A - \tfrac{1}{3}b) = kT$$

was expected to hold for three layers.

The fitting of straight lines to the curve presupposed that all the discontinuities were sharp. This treatment was balanced by drawing the F against $\log p$ curve as smoothly as possible, which led to a smoothed FA against F curve (figure 8) and placed the minimum of emphasis on the discontinuities. The FA curve indicated that the formation of a fourth layer had begun near the saturation pressure Condensation to a two-dimensional liquid was not a sufficient explanation of the changes observed with acetone By the end of the adsorption there was one acetone molecule per 10 A^2, which meant that a second layer at least must have been formed

At 50° C the condensation of the second layer did not take place until a much higher pressure was reached It was not possible to explore this region in full.

However, the errors of the last seven points from the line fitted to the first eighteen points comprised strong evidence of some phase change

The property of forming multiple layers appeared to be connected with small partially polar molecules because the changes in the isotherms of methyl and ethyl alcohols probably corresponded to the same phenomenon as found with acetone In these cases no attempt was made to smooth the results For methyl alcohol the following equations were found to fit the experimental readings:

at 25° C. $F(A-31\,3) = kT$ and $F(A-15{\cdot}7) = kT$,

and at 50° C: $F(A-28{\cdot}4) = kT$ and $F(A-14{\cdot}4) = kT$.

In both cases the co-area for the second region was almost exactly half that for the first region For ethyl alcohol at 25° C the co-areas were $22{\cdot}9 \pm 0{\cdot}3$ and $11{\cdot}7 \pm 0{\cdot}2$ A^2. Sufficiently high pressures could not be obtained to calculate the co-area for the second region at 50° C in the case of ethyl alcohol The surface pressures at the onset of the double layer adsorption were 19·8 and 28·1 for methyl, and 39·0 and 48·2 dynes/cm for ethyl alcohol at 25 and 50° C respectively These show that the tendency to form a double layer diminished with increasing temperature and chain length as expected.

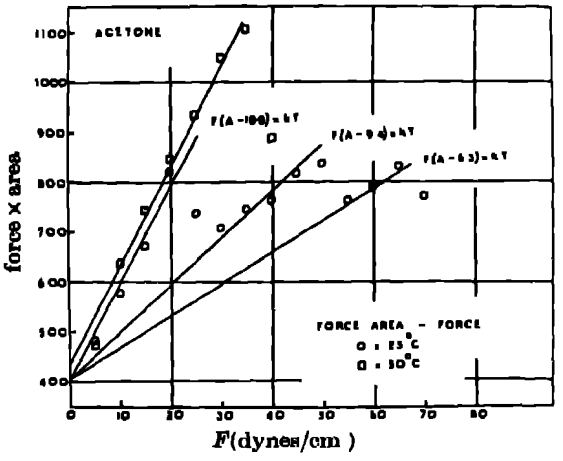


FIGURE 8

The entropy of adsorbed acetone

The extremely small value found for the entropy of adsorption is of interest Using the methods outlined in Part II (Kemball 1946*a*) the two-dimensional entropy of translation at the concentration of the standard state was found to be 29·6 e.u., compared with the three-dimensional value of 38·2 at 310·6° K and 1 atm. The minimum value that would have been expected for the entropy of adsorption was therefore 8·6, the difference between the values just quoted. However, the experi-

mental figure was only $4{\cdot}4 \pm 0{\cdot}7$ e u., which indicated that the adsorbed molecules must have possessed as much freedom to rotate and to move in two dimensions, as the vapour molecules, and further that the lost translational freedom must have been replaced by a vibration of low frequency The small energy of adsorption of 7·45 kcal./mol. was in keeping with this suggestion Two methods were used to give a rough value of the frequency of the vibration and hence the entropy associated with it The idea due to Lindemann that the amplitude of the vibrations of atoms at the melting-point is equal to the interatomic separation may be extended to calculate a 'characteristic frequency' for the acetone molecule. First, the relation (7) given by Moelwyn-Hughes (1940) was used, taking the melting-point as 174° K and the density of the solid as 0 9686, and giving the value of the frequency as 42 6 cm $^{-1}$ Secondly, use was made of the approximate relation between energy of dissociation and the frequency of a diatomic molecule. With the exception of hydrogen and deuterium the relationship is linear. Assuming that the concept could be applied to the interaction between molecules as well as to that between atoms it was found that the energy of 7·45 kcal /mol corresponded roughly to 70 cm.$^{-1}$ Both these methods were not strictly applicable to molecules but they showed agreement on the order of magnitude of the frequency. In table 9 the entropies of vibration of frequencies in that range, calculated from the relation (8) for the

$$\nu = L\nu_m^{-\frac{1}{3}} T_m^{\frac{1}{2}} M^{-\frac{1}{2}}, \tag{7}$$

where ν is the frequency, L is a constant having the value $2\,80 \times 10^{12}$, V_m is the molecular volume and T_m the melting-point

$$S_{\text{vib.}} = R\left[\frac{h\nu}{kT}(e^{-h\nu/kT} - 1)^{-1} - \ln(1 - e^{-h\nu/kT})\right], \tag{8}$$

temperature of 310·6° K, are listed. The values are of the correct order to account for the difference between the entropy of adsorption of 4 4 and the expected loss for the third degree of translational freedom, of 8·6 e u , by a slow vibration of the molecule, as a whole, in the force field of the mercury. It was presumed that the adsorption was so weak that the rotations of the molecule were hardly impaired at all. This led to a large free energy of adsorption even although the total energy was little greater.

TABLE 9 VIBRATIONAL FREQUENCIES AT 310·6° K

frequency ν cm.$^{-1}$	$x = h\nu/kT$	entropy S e u.
40	0 1846	5 34
50	0 2308	4 90
60	0·2770	4 54
70	0 3231	4·25
80	0 3693	3 98
90	0·4155	3 75
216 6	1·0000	2 05

Two-dimensional condensation

The upper regions of the adsorption isotherms for the three higher alcohols consisted of comparatively incompressible films. The F against $\log p$ curves (e.g that for n-hexyl alcohol, figure 9) were practically straight and the $\log p/F$ against F curves (e.g that for n-amyl alcohol, figure 5) were convex to the surface-pressure axis, indicating that the Volmer equation was no longer obeyed For the purpose of calculating the thermodynamic quantities for these condensed films, and for the construction of the F against A curves (figures 10, 11), straight lines were fitted to the upper parts of the F against $\log p$ curves and the compressibility of the films was ignored. The gradient of such lines gave the area occupied by each molecule in the condensed film, and the intercept on the $\log p$ axis was some measure of the free energy of adsorption

The same definition of a standard state as was used for the gaseous films was taken for these condensed films, i e. a surface pressure of 0·06084 dyne/cm. at any temperature Extrapolation of the F against $\log p$ lines was necessary as none of the films were stable down to such a low pressure The free energy of adsorption was given by

$$-\Delta G = RT \ln (p_0/p),$$

where p_0 was 1 atm and p was the pressure of vapour that would have been in equilibrium with the condensed film at a surface pressure of 0 06084 dyne/cm Hence

$$-\Delta G = 2\ 303RT \log p_0/p = 2\ 303RT(1{\cdot}8808 - \log p),$$

if p is measured in mm. of Hg. The entropy and heat of adsorption were then obtained from the temperature coefficient of the free energy The data for the condensed films of the three alcohols are summarized in table 10. As expected the heat of adsorption increased with chain length because of the greater interaction between the chains. The area per molecule remained in the neighbourhood of 20 A^2 indicating that the molecules were erect on the surface.

TABLE 10. THE CONDENSED FILMS OF THE HIGHER ALCOHOLS ON MERCURY

substance	temp. °C	area per mol A^2	ΔG cal./g.mol.	ΔH kcal /g.mol.	ΔS cal./deg g.mol.
n-butyl alcohol	25	21 8 ± 0 4	− 4200 ± 38	− 12 4 ± 1 0	− 27 4 ± 2·8
	50	23·9 ± 0·4	− 3515 ± 30		
n-amyl alcohol	25	23 4 ± 0·3	− 4868 ± 22	− 16·4 ± 0 6	− 38·5 ± 1 8
	50	21·4 ± 0 3	− 3915 ± 20		
n-hexyl alcohol	25	21 8 ± 0 3	− 5501 ± 18	− 17·2 ± 0 5	− 39 1 ± 1 2
	50	18·9 ± 0·3	− 4525 ± 12		

The force-area curves for these alcohols were calculated from the Volmer equations fitted to the first regions of the isotherms and from the area per molecule (assumed constant over the second regions) The value of the surface pressure at which two-dimensional condensation set in can be seen to diminish with chain

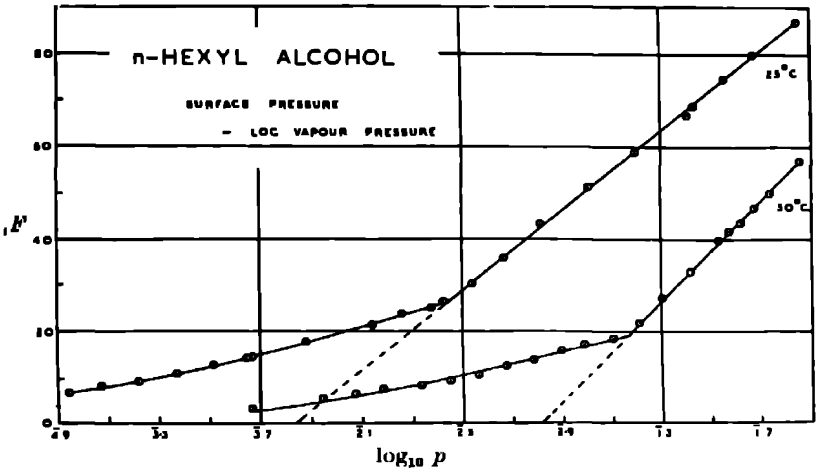


FIGURE 9

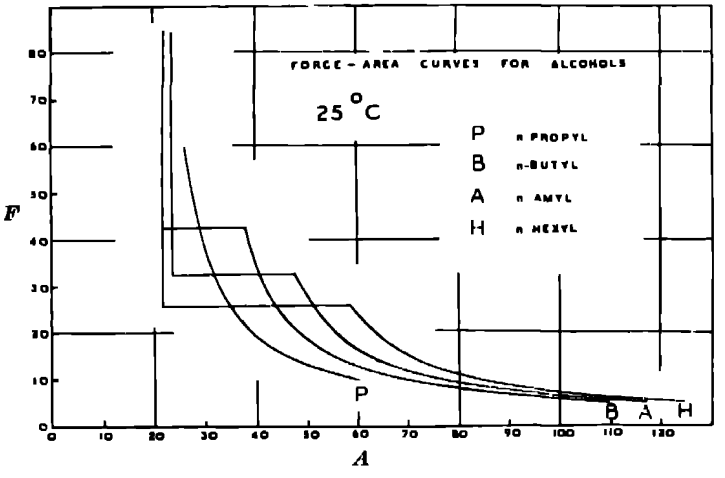


FIGURE 10

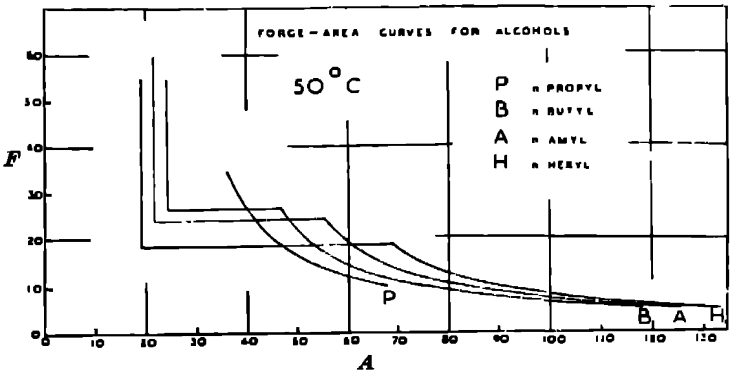


FIGURE 11

length, and the increase of the co-area of the gaseous films with chain length is also clearly shown in figures 10 and 11. A comparison of the two figures shows that for each alcohol the condensed film formed at a lower surface pressure at the higher temperature, i.e. the surface vapour pressure decreased with increasing temperature —a reversal of the normal pressure-temperature relation This interesting property was connected with the entropy change on liquefaction or alternatively with the latent heat change As the condensed film was formed from a gaseous film in which the molecules were lying in contact with the mercury there was an absorption of energy instead of an evolution. The entropy of the molecules was greater in the liquid film than in the gaseous film The effect was due to the fact that the molecules were more strongly attracted to the mercury than they were to each other, and liquefaction was only possible after the erection of the molecules.

The critical temperatures for substances on a mercury surface must be greater than the values on a water surface, because the experiments of Adam & Jessop (1926) indicated that the critical temperature of substances with 12 or 13 atoms in the chain was below room temperature No definite liquefaction was obtained for propyl alcohol on mercury but that was probably due to inability to produce a sufficiently high vapour pressure because the last few points of the $\log p/F$ against F curve (figure 4) indicated that some change had taken place in the isotherm for 25° C The large mass and high surface energy of the mercury may mean that the surface of that medium is more permanent than the surface of other liquids and give rise to a higher critical temperature for two-dimensional condensation Surface films on water generally give surface-vapour pressures of less than 0 4 dyne/cm The higher values found with mercury were due to the short chains of the substances adsorbed If the surface-vapour pressures for substances on water are plotted as $\log F$ against the number of units in the chain, an approximately linear relation is obtained, and the values for the three alcohols adsorbed on mercury also fit this relation reasonably well However, the work of Jura, Loeser, Basford & Harkins (1945) shows the liquefaction of n-heptane on silver at a surface pressure of a fraction of a dyne/cm which would not be expected from the relationship between surface-vapour pressure and chain length.

Discussion

The heats of adsorption of the different compounds showed interesting variations As had been found with the non-polar substances (Kemball & Rideal 1946) the first region of the adsorption isotherm was nearly always a gaseous monolayer obeying a Volmer type of equation (the only exception being water) The values obtained for ethyl and n-propyl alcohols probably represented the heat given out by the adsorption of an alcohol group end-on to the mercury surface, i e. about 11 kcal./g mol. The higher values obtained for methyl alcohol and water may have been partly due to the closer approach of these small molecules to the surface but were probably caused by association which was thought to occur from entropy considerations After propyl alcohol the heat of adsorption rose sharply and increased with

each extra CH_2 group in the chain, which, coupled with the increase in co-area, indicated that the molecules were lying flat on the surface The heat of adsorption of acetone was low despite the strong dipole moment The area indicated that the molecules were probably end-on to the surface and this suggests that, although the heat of adsorption might be greater if the dipole came nearer the surface, the area occupied might be uneconomically large or alternatively the freedom of the molecules might be curtailed and give rise to a lower free energy of adsorption. The low entropy of adsorption for acetone confirms that the molecules are held in some way which puts little restraint on them and allows free movement The heat of formation of the liquid films, formed with the higher alcohols, increased with chain length as expected.

The co-areas of water and methyl alcohol were much larger than might have been expected This increase may have been due to dipole repulsion between the adsorbed molecules and some trace of this still persisted with ethyl alcohol The decrease of co-area with temperature confirmed the idea of dipole repulsion, because as shown in Part I a decrease of co-area with temperature means the absorption of energy as two molecules approach to their equilibrium distance on the surface The small increase with ethyl alcohol indicated that the attractive forces were only just greater than the repulsive forces at the equilibrium separation Most of the substances adsorbed showed an increase of co-area with temperature

Contrary to the results of Cassel & Salditt (1931) water vapour was found to be adsorbed on mercury, although not strongly at 50° C The initial regions of the curves for surface pressure against vapour pressure were found to be concave to the vapour-pressure axis at all stages, whereas Cassel & Salditt had reported a convex curve with a point of inflexion for the lower alcohols This they attributed to association between the alcohol molecules, but it is doubtful if association could cause such marked effects They found that a large pressure of methyl and a somewhat smaller pressure of ethyl alcohol was required to influence the surface tension, and so on until butyl alcohol caused a lowering at extremely small pressures In all cases, once this critical pressure had been reached the effect on the surface tension of a further increase of pressure was quite marked It is possible that their mercury was covered with some non-polar contaminant, which water was not able to displace and so was not adsorbed, and the alcohols were able to displace this dirt once a certain critical pressure had been reached and this was lower the larger the non-polar part of the molecule. It is only in the cases where the attraction between the adsorbed molecules are greater than the attractions between the adsorbing surface and the adsorbed molecules, that a convex portion of the F against p curve would be expected due to association

The alcohols showed an interesting trend as regards the behaviour of the second region of the adsorption isotherm. Methyl and ethyl both formed second layers The higher alcohols from *n*-butyl upwards formed condensed films Propyl alcohol fell between the two types of behaviour. It has too long a hydrocarbon chain to allow the formation of a second layer readily—the extra CH_2 group in ethyl alcohol

delayed the formation of the double layer by several dynes/cm compared with methyl alcohol. On the other hand, the length of the propyl chain was too short to give two-dimensional condensation until a very high surface pressure was reached. One of the interesting features about the experimental results was the evidence for the discontinuous adsorption of the second layers, and in the case of acetone, of a third layer as well. Little was known about the structure of the double or triple layer thus formed, they appeared to be gaseous, obeying a Volmer equation with one-half or one-third of the normal co-area This behaviour could be explained on the concept of separate surface pressures for each layer and equal amounts in each layer This cannot be regarded as a satisfactory explanation, it is more reasonable to regard the two layers as a composite phase and not to differentiate between first- and second-layer molecules This treatment is also demanded by the application of the phase rule, for example to figure 3

Water showed no inclination to form a second layer and only a monolayer was formed at high ratios of p/p_0 The evidence obtained so far indicates that small partially polar molecules most readily form second layers It is possible that a second layer of water would not produce any lowering of surface energy. The surface tension of mercury covered with a monolayer may be regarded as the free energy of the mercury atoms and the free energy of the molecules comprising the monolayer The surface tension of mercury covered with a monolayer of water is 438 dynes/cm , whereas the corresponding values with heptane, and benzene and toluene, are 424 and 419 dynes/cm The value with water is distinctly higher and in keeping with the greater free energy of the water molecules A second layer of water will only be formed if the free energy of the mercury atoms be reduced, because it is clearly no gain towards the reduction of surface tension to replace one layer of molecules by two layers when the molecules in the second will have the same high free energy. If one supposes that one layer of water molecules is capable of saturating the field of the mercury atoms, the adsorption of a second layer would not be expected However, in the case of a partially polar molecule, it is possible a second layer might cover up some of the polar portions of the molecules in the first layer and so cause a decrease in the free energy of the surface as a whole

I have pleasure in thanking Professor Eric K Rideal, F R S , for his interest and encouragement in this work which was financed by the Ministry of Aircraft Production

References

Adam & Jessop 1926 *Proc. Roy Soc* A, **110**, 423.
Cassel & Salditt 1931 *Z phys. Chem.* A, **155**, 321
Chang 1939 *Proc. Roy. Soc.* A, **169**, 512.
Drucker & Kangro 1915 *Z phys Chem* **90**, 518
Flock, Gunnings & Holton 1931 *Bur Stand J Res* **6**, 895.
Fowler & Guggenheim 1939 *Statistical thermodynamics* Camb. Univ Press.
Giauque & Archibald 1937 *J. Amer Chem Soc* **59**, 561.

Giauque & Stout 1936 *J Amer. Chem Soc* **58**, 1144.
Gilliland & Sherwood 1934 *Industr Engng Chem.* **26**, 510.
Gordon 1934 *J. Chem Phys* **2**, 65.
Grassi 1888 *Nuovo Cim* **23**, 109
Hovorka, Lankelma & Stanford 1938 *J Amer Chem Soc.* **60**, 820.
Jura, Loeser, Basford & Harkins 1945 *J. Chem Phys* **13**, 535.
Kemball 1946*a* *Proc Roy Soc* A, **187**, 73
Kemball 1946*b* *Trans Faraday Soc.* **42**, 526
Kemball & Rideal 1946 *Proc. Roy. Soc* A, **187**, 53
Moelwyn-Hughes 1940 *Physical chemistry.* Camb Univ. Press.
Mundel 1913*a* *Z. phys Chem* **85**, 435.
Mundel 1913*b* *Z phys Chem* **85**, 475.
Pauling 1935 *J Amer Chem. Soc* **57**, 2680.
Ramsay & Young 1886 *Phil. Trans.* A, **177**, 123
Ramsay & Young 1887 *Phil Trans* A, **178**, 313
Sameshima 1918 *J Amer. Chem. Soc* **40**, 1488
Schmidt 1891 *Z. phys Chem* **8**, 628

Sublimation in a Wilson chamber

By B M. Cwilong

(*Communicated by G M B Dobson, F R S.—Received* 21 *November* 1946—*Read* 19 *June* 1947)

1 Introduction

About fifty years ago Professor C. T. R. Wilson began a classical series of experiments (Wilson 1895 to 1933) on the condensation of water vapour in both dust-free and ordinary air. The only reference to ice throughout that work is as follows (Wilson 1897*c*, p. 299)

'It is assumed here that the cloud-particles are actually liquid drops and not ice crystals, in spite of the fact that the condensation begins at temperatures much below the freezing-point* and that the temperature when the particles are fully grown is, as we shall see, also slightly below the freezing-point.'

In the present paper experiments are described in which the final temperatures after expansion are much lower These confirm that down to surprisingly low temperatures only liquid water droplets are formed, but they show that after a threshold low temperature is passed ice particles begin to be formed These ice particles are not formed by the freezing of water droplets but by the direct sublimation of water vapour to ice crystals, and the process has some essential differences from the condensation of water vapour at higher temperatures The experiments also indicate that in ordinary outdoor air there are no sublimation nuclei on which the ice crystals form when the air is supersaturated at temperatures between 0 and $-32°$ C

* About $-15°$ C.

2. Method

2 1 The initial difficulty in this work was to devise some means of detecting the presence of a few ice crystals in the cloud formed on expansion. For this purpose use was made of the well-known property of supercooled liquids to solidify, when inoculated by even the smallest particle of their solid phase. Some water* which can be easily supercooled is put in the bottom of the cloud chamber (*a* in figure 1) which consists of a glass cylinder. The whole cloud chamber is placed in a cooling bath of alcohol (*c*) contained in a Dewar flask (*d*) The cooling bath is kept at a known temperature below the freezing point. Between the cloud chamber and the piston part of the apparatus a trap (*b*), similar to liquid air traps, is introduced to prevent diffusion of water vapour from the piston part, which is at room temperature, to the cloud chamber proper.

In order to reduce turbulence after expansion, a wire-gauze partition (*D*) is placed in the tube connecting the chamber with the piston part. The piston part of the apparatus is either of the glass plunger type (*E* in figure 2) or a rubber membrane (*E* in figure 1). When only water droplets are formed on expansion the supercooled water remains liquid, but if one ice crystal is formed it must eventually fall into the water or on to the glass walls which are also wet with supercooled water so that the whole will freeze.

In the case of dust-free air the threshold temperature, below which the first ice crystals are formed, is definite to within about 0·1° C. To measure it accurately a special cloud chamber was made consisting of three legs, all connecting with the same expansion chamber, but cooled in three different baths to slightly different temperatures (figure 2) Since the expansion ratio is necessarily the same in all three legs, the final temperatures after expansion will differ from each other as the bath temperatures differ The bath temperatures were so chosen that ice particles were formed in one or two legs but not in all, so that upper and lower limits to the temperature were obtained

In some cases, especially when very low final temperatures were investigated and where the bath temperatures were necessarily too low to allow the use of supercooled water, the presence of ice crystals was observed visually, but this second method is necessarily less accurate than the first. In the second method the water on the bottom and sides of the cloud chamber was frozen and a clear window made by evaporating the ice from a small area near the bottom of the cloud chamber which consisted, in this case, of a glass cylinder 70 cm long. If a few ice crystals were formed they would fall through the cloud of supercooled water droplets (i.e through air supersaturated with respect to ice) and grow so rapidly that by the

* It was found that easily supercooled samples of water and heavy water could be obtained by repeatedly distilling the water in a vacuum without bubbling and afterwards boiling the sample until the greater part was evaporated Water prepared in this way can be supercooled as much as 10° C The lowest supercooling obtained by this method was 23° C and different samples each have their own freezing point which they appear to keep indefinitely

time they reached the bottom of the chamber they could be seen by eye, glittering in the light of an arc lamp

2·2 *Calculation of the adiabatic cooling.* By the above methods the expansion ratio which, with any given bath temperature, will just result in some ice particles being formed can be fixed with great accuracy but it cannot be assumed that the final temperature immediately after expansion can be obtained by applying the usual adiabatic formula. In fact, the temperature change in our case was generally

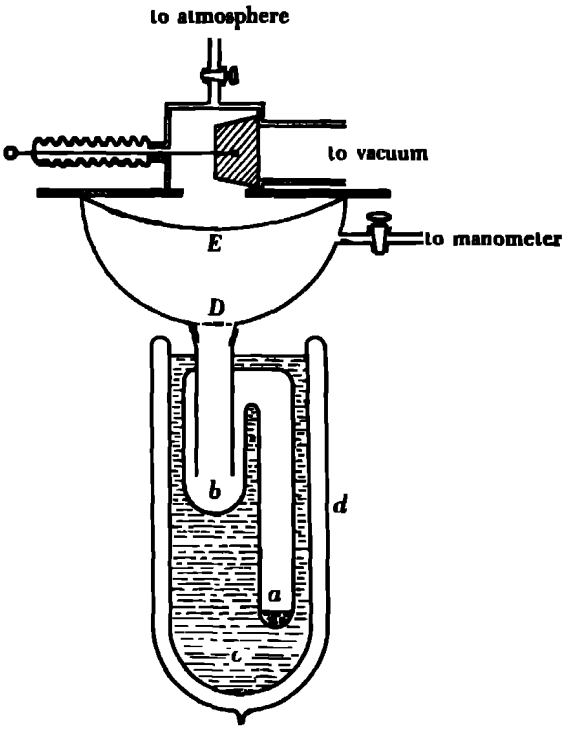


FIGURE 1

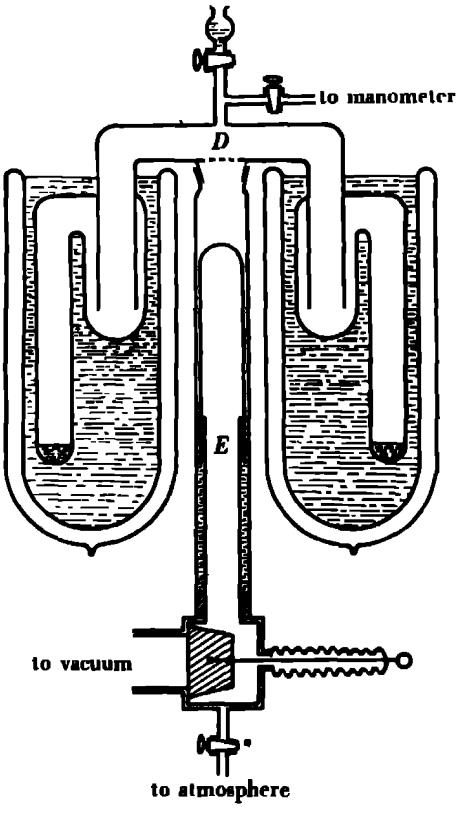


FIGURE 2

some 15 % less than this. The fact is often overlooked that the ratio of the specific heats (γ) without experimental correction can only be applied in calculations of expansion cooling to very big chambers (over about 50 l). To evaluate the effective value of γ, room air of measured humidity was specially enriched in condensation nuclei by passing it over a glowing wire This air was then admitted into the dry expansion chamber and a series of expansions made, with successively increasing expansion ratio until a fog was just formed, indicating that the air had been cooled to its dew-point When the expansion chamber is illuminated by the beam of the arc light this 'fog-point' can easily be determined within 0·2° C.

3. Results

3 1. *Sublimation in dust-free air* When the air in a cloud chamber is first freed from all dust nuclei by several expansions not exceeding a ratio of 1·25 and then the chamber is placed in the cooling bath at a temperature between -1 and $-10°$ C, consecutive expansions, increasing in ratio, show the following phenomena (1) up to a ratio of about 1·25 no particles at all can be observed and the supercooled water remains liquid; (2) for expansion ratios between 1 25 and 1 38 particles are formed as described by C. T. R. Wilson and the supercooled water remains liquid, indicating that only water droplets are formed, (3) as soon as the temperature at the end of the expansion falls to -41 2° C the supercooled water usually freezes although there may be in this case a delay of from 1 sec. to a few minutes during which time ice crystals are falling, (4) when the final temperature is a few tenths of a degree lower still, freezing always takes place immediately, indicating that many ice particles are present and inoculation is not delayed

When the initial temperature of the chamber is below $-10°$ C (with ice covering walls and at the bottom) and the expansion ratio is such that the final temperature is below -41 2° C, several ice particles glittering in the light are observed even when the expansion ratio is smaller than 1·25 They were certainly observed with the expansion ratio as low as 1·07 and this value is probably rather the limit of accurate observation, than of the process It should be noted that this is in air freed from all ordinary nuclei and in which no condensation would occur if the final temperature had been above $-41·2°$ C This shows explicitly that the ice particles are formed by direct sublimation from the vapour

When the final temperature reaches about $-70°$ C a change in appearance of the cloud is observed During the first few seconds after expansion, instead of the fine crystalline dust found at higher temperatures a small shower of quite big grains of ice falls. After this follows a cloud of crystalline dust which takes several minutes to settle.

At final temperatures of about $-100°$ C only the grain shower remains Even when the expansion ratio greatly exceeds Wilson's cloud condensation ratio no other particles are visible Below about $-120°$ C no particles are observed

When the chamber was filled with CO_2 from industrial high pressure cylinders no appreciable difference of temperature was observed from the results with air

If the air in the cloud chamber is ionized by exposure to X-rays before the expansion, many more ice particles are formed, but there is no change in the critical temperature of -41 2° C

3·2. *Sublimation in outdoor air.* A series of experiments with outdoor air were made in the three-leg chamber on the flat roof of the Clarendon Laboratory, Oxford, and during very clear conditions at Shotover, 4 miles outside Oxford The chambers were well ventilated until 3 min. before each expansion It was found that the maximum temperature for the first appearance of ice particles ($-41·2°$ C for clear air) was now changed to $-32·2°$ C. The threshold temperature is not so definite as

with dust-free air but varies from one experiment to another with a range of about 1° C. The place, season of the year and weather have no appreciable influence on this temperature.

3·3 *Sublimation in contaminated air* When the air in the chamber was artificially contaminated before each expansion by introducing tobacco smoke, it was then found that the maximum temperature for the appearance of ice is much higher and about $-27°$ C ($\pm 0{\cdot}5°$ C).

3·4. *Sublimation with heavy water in dust-free air* It was found with a sample of heavy water *d*.20/4.1 10436 that the highest temperature of the appearance of ice was $-37{\cdot}4°$ C

4. Remarks on experimental data

The experimental data are conveniently summarized in the form of a table at the end of this paper.

4·1. *Ice in dust-free air* From experiments in dust-free air it can be concluded that there are no centres on which either sublimation or freezing of water droplets can start at a temperature above $-41{\cdot}2°$ C Below $-41{\cdot}2°$ C very small expansions always give a scanty shower of ice crystals C T R Wilson has shown that in dust-free air negative ions are the most active centres of condensation As the density of the shower of ice particles in our case is of the order of Wilson's 'rain-like' condensation, this seems to indicate that clusters of ions and single ions are acting as genuine nuclei of sublimation too This fact is also confirmed by X-ray evidence: when the expansion chamber is illuminated by X-rays the density of the shower increases. In the case of sublimation the action of centres of sublimation is quite different from Wilson's condensation nuclei because it seems only to depend on the final temperature and not on the degree of supersaturation. The following argument indicates that in the first method (using supercooled water at the bottom of the chamber as a detector and consequently large expansion ratios) the initial ice particles must have been formed by direct sublimation and not by the freezing of water drops In the second method we can be certain that only ice particles are formed because the expansion ratio was too small to allow of any drop formation From this and the coincidence of the threshold temperatures found by the two methods it is legitimate to argue that ice particles appearing in the mixed cloud of water and ice encountered in the first method, are formed by direct sublimation and not by the freezing of water drops.

Are ions the only centres on which sublimation takes place? The mixed fogs of water drops and ice particles are observed down to $-70°$ C But the number of ice particles increases as the final temperature after expansion decreases. In large expansions when the final temperature is below $-60°$ C, an ice cloud is formed which in density resembles Wilson's cloud-like condensation rather than his shower-like condensation. For the formation of such a cloud the number of ions in the

chamber is not sufficient so that many ice particles are formed either by subsequent freezing of water droplets or by sublimation on neutral centres. The majority of water drops in mixed fogs evaporate quickly in the presence of ice particles, as indicated by the visible decrease in density of the mixed cloud in the first minutes after expansion. (There is no similar change when purely water clouds are formed.) But such decrease becomes less and less marked when approaching $-60°$ C, and the ice cloud shows no change in density at temperatures about $-65°$ C. Some water droplets are probably inoculated by ice particles owing to different velocities of fall. It is also possible that in the mixed cloud some water drops, being near to ice particles, will evaporate quickly and lose enough heat to freeze. A similar occurrence was observed during the earlier work on sublimation on solid surfaces: supercooled water drops were deposited on a gold surface from air with a dew-point of $-30°$ C The air supply then was suddenly changed to much drier air, with the dew-point $-65°$ C The drops started to evaporate quickly, the majority freezing at once.

4·2. *Ice in outdoor air and in contaminated air* The process of the appearance of ice in this case is different. More experimental evidence is necessary for answering the main questions are there any other nuclei acting as genuine sublimation centres, or is it just the freezing of water drops?

An interesting problem from the meteorological point of view is to find the threshold temperatures at which sublimation commences at various cloud levels

TABLE SUMMARIZING ALL EXPERIMENTAL RESULTS

air	expansion ratio (E)	temperature (t) initial	temperature (t) after expansion	products of expansion
		C. T. R. Wilson (1897)		
outdoor	$E>1$	about room	—	dense fog of water droplets
clean	$E<1·25$	,,	—	none
	$1·25<E<1·38$	,,	—	shower of water droplets
	$E>1·38$	,,	$t>-15°$ C	dense fog of water droplets
		B. M. Cwilong (1944)		
outdoor	$E>1$	$t<0°$ C	$t>-32·2°$ C	dense fog of water droplets
			$t\leqslant-32·2°$ C	dense fog of water droplets and ice particles
clean	$E<1·25$	,,	$t>-41·2°$ C	none
			$t\leqslant-41·2°$ C	shower of ice particles
	$1·25<E<1·38$	,,	$t>-41·2°$ C	shower of water droplets
			$t\leqslant-41·2°$ C	shower of water droplets and ice particles
	$E>1·38$	$-10°C<t<0°C$	$t>-41·2°$ C	dense fog of water droplets
			$t\leqslant-41·2°$ C	dense fog of water droplets and ice particles
clean and outdoor	—	about $-30°$ C	$-100°C<t<-70°$ C	shower of soft hail and fog of ice dust
		,,	t about $-100°$ C	shower of soft hail
		,,	$t<-120°$ C	none

The work described above was carried out for the Meteorological Office, and I am indebted to the Director, Sir Nelson Johnson, and to the Professor of Experimental Philosophy in Oxford University, Lord Cherwell, F R S., for enabling me to pursue this research at the Clarendon Laboratory, Oxford It is a pleasure to acknowledge my indebtedness to the Reader in Meteorology, Professor G. M B Dobson, F.R S , for his keen interest in my research and many stimulating discussions and suggestions

References

Cwilong, B M 1945 *Nature*, **155**, 361.
Partington, J. R & Howe, A. B. 1924 *Proc Roy Soc* A, **105**, 225.
Powell, C. F. 1928 *Proc Roy Soc* A, **119**, 553
Wilson, C T. R 1895 *Proc Camb. Phil. Soc.* **8**, 306
Wilson, C T R 1897*a* *Proc Camb. Phil Soc* **9**, 333, 392
Wilson, C T R. 1897*b* *Proc Roy Soc* **61**, 240
Wilson, C. T. R 1897*c* *Phil Trans* A, **189**, 265
Wilson, C T R. 1899 *Phil Trans* A, **192**, 403
Wilson, C. T. R. 1900 *Phil Trans.* A **193**, 289.
Wilson, C T. R 1933 *Proc Roy Soc* A, **142**, 88

Studies on silk proteins

I. The properties and constitution of fibroin. The conversion of fibroin into a water-soluble form and its bearing on the phenomenon of denaturation

By D Coleman and F O Howitt

The British Cotton Industry Research Association (Silk Section), Shirley Institute, Manchester

(*Communicated by Sir Robert Pickard, F R S* —
Received 13 *February* 1945—*Revised* 21 *October* 1946)

[Plate 5]

Some of the constituent amino-acids of fibroin (degummed silk) are determined Special attention is directed to histidine, owing to its use in the calculation of the molecular weight of fibroin A value of 0 45 % has been found by methods in which the histidine is isolated as nitranilate or di-(3 4 dichlorobenzenesulphonate) Other values obtained are serine 12 0 %, threonine 1 5 %, tyrosine 10 6 %, and proline 1 5 % Hydroxyproline appears to be absent, but the presence of small amounts of some hydroxyamino-acid other than serine and threonine is indicated

The mean residue weight of fibroin is determined by three methods, one of which is a new method based on analysis of the complex formed between fibroin and cupri-ethylenediamine This method gives a Cu fibroin-N ratio of 1 1 02 and, if allowance is made for co-ordination with the tyrosine hydroxyl group, an equivalence of 1 964 atoms of peptide-nitrogen to 1 atom of copper is obtained. The three methods give an average value of 78 0 for the mean residue weight of fibroin. This value, together with the most acceptable data for amino-acid constituents, indicate that the unidentified anhydro-residues (about 20 %) have a mean residue weight of about 107

Evidence is presented that fibroin contains no amide-nitrogen Methods for the determination of amide-nitrogen at present in use, which indicate a content of 1 to 2 %, are considered to be unreliable

Fibroin dissolved in cupri ethylenediamine gives, on neutralization and dialysis of the resulting solution, a water-soluble protein The production of this water-soluble protein is attended by little or no degradation of the original fibroin as shown by determinations of fluidity, amino nitrogen, and acid- and alkali-combining power

The water-soluble protein is precipitated by the normal protein-precipitating reagents, but in every instance examined the precipitated material exhibits an insolubility comparable with that of the original fibroin.

Factors responsible for the solubilization process are investigated and data for molecular weight, titration values, ultra-violet absorption spectra, reducing activity, optical rotation, tryptic hydrolysis, and viscosity for both soluble and dispersed fibroin are given Soluble fibroin has $[\alpha]_D^{18} - 53\ 1°$ and dispersed fibroin $[\alpha]_D^{18} - 58\ 9°$, both in aqueous media

The preparation and properties of films and filaments of fibroin are described Films of fibroin can be prepared that are water-soluble On stretching, these films show strain-birefringence, acquire considerable tensile strength, and become insoluble in water, but X-ray examination gives the β-keratin pattern for both the stretched and unstretched films

Reasons are advanced for considering the water soluble form of fibroin to be the native or renatured protein and the original protein to be the denatured form

The denaturation of fibroin is discussed on the basis that denaturation is essentially an unfolding of a coiled long-chain molecule The subsequent aggregation of the uncoiled molecules to give an insoluble product is considered to be a secondary process Some aspects of protein and polypeptide chains as macro-molecules are also discussed

1. Introduction

The thread extruded by the silkworm, *Bombyx mori*, consists essentially of two proteins, one of which (fibroin) forms the substance of the two component filaments and the other (sericin) a cementing material coating the two filaments and joining them together The sericin layer is readily removed by mild alkalis, e g by boiling soap solution, leaving the filaments of fibroin in a fairly pure condition Subsequent treatment of the fibrous protein with dilute acetic acid, followed by extensive washing with distilled water, yields a protein material of high purity. This ease of preparation probably accounts for the early attention that fibroin received (mainly by Fischer and his pupils) and for the extensive data that have been published for its constitution.

The great disadvantage attending work on fibroin in the past had been its insolubility. even in dilute acids or alkalis. It is possible, however, to convert this insoluble fibrous protein into a form that is freely soluble in water, which considerably broadens the field of investigation.

The problem of the structure of fibroin is inseparably bound up with its amino-acid composition, especially in the light of the Bergmann-Niemann theory of periodicity of amino-acid distribution. Fibroin, with its freedom from cystine and methionine, has a relatively simple amino-acid composition, and an extension of the available analytical data (see Synge 1943) permits the advancement of a structural theory that is compatible with the composition of the fibroin molecule and of a possible explanation of the characteristic properties that are a reflexion of its biological role.

2 Experimental

A *Determination of some amino-acids in fibroin*

(i) *Histidine* The histidine content of 0 07 % found by Vickery & Block (1931) has been generally accepted for fibroin. An examination of the analytical method used by these workers, however, indicated that this value is probably too low The method first adopted was a modification of the combined procedures described by Block (1934, 1940), the histidine being isolated as nitranilate This method gave an actual yield of 0 21 %, and evidence was obtained that absorption losses (due to the use of baryta as hydrolyzing agent) had occurred, indicating that the true content was higher than this value Subsequently, a modification of the method of Vickery & Winternitz (1944) was applied to hydrochloric-acid hydrolysates of fibroin, and this gave yields of histidine [isolated as di-(3 4-dichlorobenzene-sulphonate) or nitranilate] of the order of 0 30 to 0·35 % During the development of this latter method, it was found that the silver salt of histidine has an appreciable solubility (of the order of 0 001 %) in the aqueous media used in precipitation and washing and, if allowance is made for this solubility, final values of 0·40 to 0 45 % are obtained.*

* Since this paper was first submitted for publication, Dunn, Camien, Shankman & Rockland (1945) have found a histidine content of 0 34 % and Stokes, Gunness, Dwyer & Caswell (1945) a content of 0·41 %. Both these groups of workers used a micro-biological method.

(ii) *Serine and threonine* The presence of relatively large amounts of serine in fibroin was not suspected until 1941 when Nicolet & Saidel (1941), following up some work on the oxidation of hydroxy-acids with periodic acid, utilized this reagent for the determination of serine and threonine in a silk fibroin hydrolysate They found 13 6 % of serine and 1 28 % of threonine Subsequent determinations by Bergmann (1941), Martin & Synge (1941), and Gordon, Martin & Synge (1943), do not show satisfactory agreement In the present work, the method described by Nicolet was used for serine and threonine and, in addition, the total hydroxy-amino-N was determined by titration of the ammonia formed in the periodate oxidation of the hydrolysate The mean values obtained for serine and threonine were 12·60 and 1 51 %, respectively; these values correspond to a hydroxyamino-acid-N of 10·13 % of the total N, and the found value of 11 1 % may indicate the presence of small amounts of some other hydroxyamino-acid, e g *β*-hydroxy-glutamic.

(iii) *Tyrosine* There is also some divergence in the literature regarding the content of tyrosine in fibroin (Meyer, Fuld & Klemm 1940) Application of the colorimetric method of Zuverkalov (1927) gave a mean value of 10 60 %

(iv) *Proline* Attempts to separate proline as the alcohol-soluble copper salt were unsuccessful The colorimetric method of Guest (1939) gave reproducible results, and a mean value of 1·50 % for fibroin hydrolyzed by baryta was obtained Hydroxyproline was not found by this method

B *Determination of mean residue weight of fibroin*

The mean residue weight (M R W) of some proteins can be calculated with fair accuracy from available X-ray and density data With fibroin, however, no definite value has been assigned to the side-chain spacing as found by X-rays Kratky & Kuriyama (1931) deduced a spacing of 4·5 to 6 1 A, and from this it follows that, for a density of 1·4, the M R W is between 62 and 84 The method often adopted is that of calculating the value from the number of gram-residues of each amino-acid in 100 g of the protein, the accuracy being dependent on the availability and accuracy of the analytical data (cf Astbury 1934) Bergmann & Niemann (1938) used this method, but their value of 84 for the M.R W is a miscalculation, the data used giving 70·1 The method of Chibnall (1942) is an improvement, the value being calculated from the N content after making allowances for the non-peptide-N in the molecule Thus with fibroin (100 g), the N content of which was found to be 18·35 % in agreement with Colombo (1932), the non-peptide-N (due to arginine, lysine and histidine) is 0·395 g (see table 2), and hence the peptide-N is 18 350 − 0 395 = 17 955 g. and the M R W is 1400/17 955 = 77 96 Yet another method of determining the M R.W is by hydrolyzing a known weight of the protein and weighing the amino-acids obtained (Chibnall 1942) A typical analysis of fibroin by this method is given in table 1 A value of 80 was obtained as the mean of several determinations that varied from 76 to 84 No greater accuracy could be expected,

since the hydrolysis yields appreciable amounts of products other than amino-acids and also any error in the total N is multiplied five times in the final result.

TABLE 1 DETERMINATION OF M.R W. OF FIBROIN

total N	128 4 mg ≡ 0 700 g. of fibroin
amino-N	120·7 mg
ammonia-N	3 4 mg
HCl	0·2826 g.
amino-acid hydrochlorides	1 1396 g
amino-acids (by diff)	0 857 g

$$\frac{\text{M R W}}{\text{M R.W} + 18} = \frac{0\ 700}{0\ 857}, \quad \text{whence M R W} = 80\ 2$$

A new method for the determination of M R W. depends on analysis of the complex formed between fibroin and cupri-ethylenediamine (cf Takamatsu 1933) Fibroin dissolves readily in cupri-ethylenediamine (prepared by dissolving cupric hydroxide in aqueous ethylenediamine),* and the resulting solution yields on addition of alcohol a precipitate of the complex The question of complex formation between copper and proteins is of some importance Jesserer & Lieben (1937) found that alkaline copper-protein solutions had a saturation value of Cu N ratio 1.6, but subsequently Kuntzel & Droscher (1940) obtained a ratio of 1 2 38 for gelatin and normal saturation ratios of 1 6 9 for caseinogen and 1.4 1 for ovalbumin, the latter gave unstable saturation systems with a ratio as low as 1 1 25 It was found that fibroin yielded a complex of a more constant composition than is indicated by these observations Experiments on the effect of concentration of En and $Cu(OH)_2$ on the solubility of fibroin showed that (i) with a constant concentration of $Cu(OH)_2$ of 0 10 g mol /l , an increase in concentration of En of from 0 216 to 1·290 g mol /l produced a steady diminution in solubility of fibroin and (ii) successive dippings of samples of fibroin (in the form of degummed silk cloth) in a solution containing 0 15 g mol of $Cu(OH)_2$ and 0 180 g mol of En per l resulted in the removal of only 50 % of the En from the solution, whilst, if excess of $Cu(OH)_2$ were present, En was completely removed from the solution It follows that the mechanism of the dissolution of fibroin in cupri-ethylenediamine may be represented by the equations·

$$CuEn_2(OH)_2 \rightleftharpoons CuEn(OH)_2 + En,$$

$$F + CuEn(OH)_2 \rightleftharpoons CuEnF(OH)_2,$$

where F is the combining equivalent of fibroin. For the determination of M R W., fibroin (2 g) was dissolved in 6/8 CuEn (100 ml); approximately 5 vol of methanol were added and the resulting viscous precipitate was separated by decantation and washed with methanol until the washings were colourless The residue was then dissolved in water and the solution made up to 100 ml. Aliquots were analyzed for

* This reagent will be referred to as x/y CuEn, where x and y are the number of grams of $Cu(OH)_2$ and ethylenediamine, respectively, in 100 ml of the solution.

total N, Cu (wet digestion and iodometric titration), amino-N (Van Slyke), and fibroin-N (acidification and precipitation of fibroin by acetone) The mean molecular ratio of Cu En fibroin-N obtained (9 determinations) was 1 0 98 1 92 The interpretation of the Cu fibroin-N ratio of 1 1 92 necessitates some consideration The alkalinity of the CuEn reagent should result in combination of the CuEn with the tyrosine hydroxyl groups, and this supposition was supported by the following experiments

(i) Carvacrol (2 ml) was mixed with 12/16 CuEn (10 ml) at room temperature The mixture formed an almost solid crystalline mass with evolution of heat After cooling at 5° C, it was pressed out on a porous plate and washed on the plate with minimum amounts of methanol followed by anhydrous ether The product was *cupri-ethylenediamine dicarvacrol*, no definite melting-point, readily decomposing on heating (N 12 85 %, Cu 11·45 %, calc for $CuEn_2[O\ C_6H_3(Me)(CHMe_2)]_2$, N 13 2 %, Cu 11 55 %).

(ii) Tyrosine ethyl ester hydrochloride, m p. 104 to 105° C (uncorr) (N 6 85 %, calc 6·70 %), prepared by the method of Fischer (1901), was ground up in a mortar with acetic anhydride (1 26 mol) for 10 min Water was added and, after mixing, the mixture was filtered The solid residue was crystallized twice from ethanol to give N-*acetyltyrosine ethyl ester*, m p 77 5 to 78 5° C (uncorr) (N 5 53 %, calc for $HO\ C_6H_4\ CH_2\ CH(NHAc)\ CO_2Et$, 5 58 %) The compound was soluble in dilute NaOH but insoluble in dilute Na_2CO_3, indicating the phenolic hydroxyl group to be free Attempts to prepare the $CuEn_2$ derivative failed, but the solubility of the ester (1 ml of 6/8 CuEn dissolved 0 47 g) indicated the formation of cupri-ethylenediamine di-(*N*-acetyltyrosine ethyl ester) (theoretically 0 484 g).

Hence in the relation that 1 g -atom of Cu combines with $1\ 92 \times 5\ 45 \times 14 = 147$ g of fibroin, allowance must be made for the number of g -atoms of Cu combined with the tyrosine in 147 g of fibroin, viz $10\ 6/362 \times 147/100 = 0{\cdot}043$ This means that 0·957 g -atoms of Cu combine (via the peptide-N) with 147 g of fibroin, and the M R W is therefore $147/2 \times 0\ 957 = 76\ 8$. It will also be observed from the data for distribution of N in fibroin (table 2) that for 100 g of fibroin there are 1 3107 – 1 2820 g equiv. of peptide-N and 1 2820 + 0 0586 = 1 3406 g equiv. of peptide-N + tyrosine hydroxyls, providing it is assumed that the unknown amino-acids contain no reactive N apart from a single amino-N. It follows that each g -atom of Cu combines with $1\ 92 \times 1\ 3406/1\ 3107 = 1\ 964$ g equiv of peptide-N, and this justifies a picture of copper atoms each being linked to two atoms of peptide-N along practically the whole length of the protein chain.

From the results of the above three methods, a M R W of 78 may be suggested as a fairly accurate value for fibroin The availability of this value makes it possible to construct a table (table 2) of the amino-acid distribution in fibroin that is more complete than that given by Bergmann & Niemann (1938) and somewhat more accurate than that given by Synge (1943), the table is based on the amino-acid figures given above and on the most recently determined figures for the other principal amino-acids, viz glycine and leucine (Moore & Stein 1943), alanine

(Gordon *et al.* 1943), valine and arginine (McMahan & Snell 1944). The histidine content of fibroin has been taken as 0 47 % so as to be compatible with the molecular weight of fibroin (described later in this paper), similarly, the value of 0·25 % for lysine given by Vickery & Block (1931) has been changed to 0 44 % No allowance has been made for amide-N, for the experimental work (described below) indicated that the content of amide-N in fibroin is either nil or negligible It will be seen from table 2 that 20 % of the protein is still unidentified. From the M R W of 78·0, the number of g equiv. of anhydro-residue in 100 g of fibroin is 1·2820; hence the corresponding number unidentified is 0·1926 There being 20 56 g of unidentified anhydro-residues, it follows that the M R W of this fraction is 107, i e the mean molecular weight is 125. The presence of phenylalanine and aspartic and glutamic acids has been reported by Fischer and others and the mean molecular weight of 125 for the missing amino-acids is probably due to these acids, and, possibly, to one or more of the three isomeric leucines (mol. wt 131).

TABLE 2 AMINO-ACID COMPOSITION OF SILK FIBROIN

(1)	(2)	(3)	(4)	(5)	(6)
amino-acid	g of amino-acid from 100 g. of fibroin	g of anhydro-residue in 100 g of fibroin	g equiv of anhydro-residue in 100 g. of fibroin	g.-atoms of N in 100 g. of fibroin	frequencies of the residues
glycine	42 3	32 15	0 5640	0 5640	2·27
alanine	24 5	19 55	0 2753	0 2753	4 66
serine	12 6	10 44	0·1200	0 1200	10 68
tyrosine	10 6	9 55	0·0586	0 0586	21 88
valine	3 2	2 78	0 0274	0 0274	46 80
proline	1 5	1 27	0·0130	0 0130	98 61
threonine	1·5	1 27	0 0126	0 0126	101 70
leucine	0 8	0 69	0 0061	0 0061	210·0
arginine	1 05	0 94	0 0064	0 0256	200 3
lysine	0 44	0 39	0 0030	0 0060	427·0
histidine	0 47	0 41	0 0030	0·0090	427 0
	total	79 44	1 0894	1·1176	
	unidentified	20 56	0 1926	0 1931	
		100 00	1 2820	1 3107	

C. *Amide-nitrogen content of fibroin*

The production of ammonia by hydrolysis of proteins with acids is usually ascribed to the presence of amide-N linkages in the molecule The first serious consideration of this type of nitrogen was due to Osborne, Leavenworth & Brautlecht (1908), who pointed out that the ammonia would be quantitatively accounted for if it were assumed to arise from hydrolysis of amide groupings formed from carboxyl groups of the dicarboxylic acid residues More direct evidence for the presence of amide groups was provided by Damodaran (1932) and by Damodaran, Jaaback & Chibnall (1932) who isolated asparagine from enzymic hydrolysates

of edestin and gliadin Methods at present in use for the determination of amide-N are mainly based on the controlled hydrolysis of the protein by mineral acids and, owing to the importance that has recently been attached to the significance of amide groups in protein structure, it is desirable to examine these methods somewhat critically.

A preliminary indication of the amide-N of fibroin is afforded by an examination of its hydrolysis by 0·1 N-NaOH, a reagent that may be assumed to decompose amide groups with liberation of the equivalent amount of ammonia. Using an ordinary Kjeldahl distillation apparatus, 1·0 g. of dry fibroin was boiled with 200 ml of 0·1 N-NaOH and the evolved ammonia absorbed in 25 ml of 2 % boric acid. The receiver was changed at intervals, the contents being made up to 250 ml

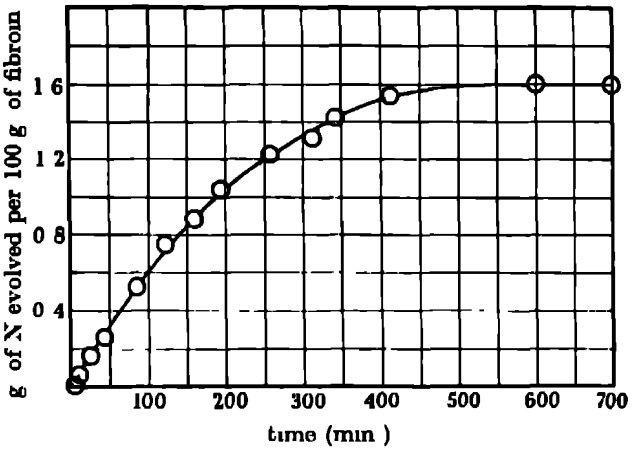


FIGURE 1. Rate of formation of ammonia from fibroin (1 g.) and boiling 0·1 N-NaOH (200 ml.)

and aliquots titrated with 0·01 N-HCl It was necessary to run water into the distilling flask at the same rate as the distillation proceeded, in order to maintain the alkali at the same concentration The rate of liberation of ammonia is shown by the curve in figure 1 This curve shows that a maximum of 9 % of the total nitrogen is obtained as ammonia, a value which corresponds closely to the serine content of fibroin In this connexion, Nicolet & Shinn (1941) refluxed silk fibroin with 0·1 N-NaOH for 2 hr and found the serine + threonine content to decrease by 0·556 m equiv./g., corresponding to an ammonia loss of 0·778 g. of nitrogen per 100 g. of fibroin, a value that falls on the curve in figure 1. These results, together with the absence of an inflexion in the curve of figure 1, suggest that there is no amide-N in fibroin, for it is very unlikely that any true amide-N in the molecule would resist the hydrolysis for 2 hr According to Chibnall (1942), however, the amide groups may be linked to other carboxyl groups to give linkages of the type $-(CH_2)_2-CO-NH-CO-CHR-$, but Bergell & Feigl (1908) have found that diglycinimide, $NH_2-CH_2-CO-NH-CO-CH_2-NH_2$, is readily hydrolyzed by

alkali to the acid $NH(CH_2COOH)_2$, whilst we have found that phthalimide is hydrolyzed to phthalic acid on boiling with 0·1 N-NaOH (50 mg of phthalimide in 200 ml of the alkali is decomposed to the extent of 24 % in 2 hr) These observations indicate that fibroin is essentially free from amide groups

The amount of ammonia-N formed on digesting fibroin with acids depends, as with most proteins, on the conditions employed Hydrolysis by boiling 8N-sulphuric acid for 16 to 20 hr gives 2 to 3 % of the total nitrogen as ammonia-N Chibnall (1942) records a value of 2·2 %, but does not state the method employed. Gordon *et al.* (1941) recommend hydrolysis by 8N-hydrochloric acid at 37° C, and, in our hands, this method has given values of 0 92 % (of the total N) after 4 days increasing steadily to 1 33 % after 20 days of hydrolysis This ammonia-N does not represent amide-N of the fibroin molecule It appears advisable to postpone unqualified acceptance of the published data for the amide-N of proteins until a more reliable method of determination has been established

D *Preparation and properties of dispersed and water-soluble forms of fibroin*

In order to determine the extent of chemical degradation in degummed silk threads, a fluidity test along similar lines to that of Clibbens *et al* (1928, 1936) for cellulose was devised by Garrett & Howitt (1941) During the development of this test, it became apparent that the fibroin could be converted into a water-soluble form For the test, the degummed silk is dissolved in 6/8 CuEn to give a solution 10 % (w/v) in fibroin and, after exactly 3 min from the commencement of dissolution, the solution is neutralized with an equal volume of 1 25N-acetic acid, the resulting solution (pH approx 8) being that used for viscometric examination Dialysis of the neutralized solution results in removal of cupri-ethylenediamine diacetate and, if the final stages of dialysis are carried out in the presence of KCN, a slightly opalescent, copper-free protein solution is produced If the dialysis is continued for about 2 weeks, part of the fibroin precipitates in the form of microscopic fibrils about 10μ in length (figure 10, plate 5) Usually, however, when the dialysis is not unduly prolonged, the solution remains fairly clear for about 14 days and then suddenly sets to a gel One outstanding property of the water-soluble fibroin is its high surface activity Vigorous stirring removes the whole of the protein from solution in an hour or less When poured from one vessel to another, the solution leaves solid fibroin in the form of fibres on the sides of the vessels There is invariably a solid film on the surface of the liquid, which on removal to a microscope slide has the appearance shown in figure 11, plate 5

The protein material in aqueous solution could be completely precipitated by the usual protein-precipitating reagents, e g trichloroacetic acid, alcohol, acetone, the material thus precipitated being completely insoluble in water It was first established that all of the fibroin (determined as nitrogen) could be recovered in the water-soluble form, excepting for 2 to 3 % that was obviously due to small amounts of material that became insoluble and adhered to the inside of the dialysis



FIGURE 10



FIGURE 12



FIGURE 11

FIGURE 10 Rod-like fibrils (× 1150) that separate on dialysis of 50 50 fibrom mixture

FIGURE 11 Fibrils (× 176) in surface film from solution of fibrom

FIGURE 12 Material that separated during attempted crystallization of fibromogen from glycerophosphate buffer (× 1150)

(*Facing p.* 152)

shell. In no instance could protein- or peptide-N be detected in concentrates of the liquid from outside of the dialysis shell. It was then necessary to be sure that no or little degradation of the protein molecule occurs during the solubilization process The degree of degradation was followed by means of the change in (i) fluidity, (ii) amino-N, and (iii) titration data

(i) *Fluidity* Degummed silk (5 g.) was dissolved in 100 ml of 6/8 solvent, the solution was neutralized after 3 min and, after determination of fluidity, the protein was precipitated by pouring the solution, with stirring, into six times its volume of acetone The powder thus produced was washed several times with acetone followed by ether and finally dried in a vacuum desiccator This dissolution and reprecipitation was twice repeated The increases in fluidity (in reciprocal poises), viz. 5 5→7 4→8 6→9 5, corresponded to molecular weights (interpolated from osmotic pressure—fluidity data, to which reference is made later) of 32,800, 32,100, 31,200 and 30,700 respectively

(ii) *Amino*-N Aqueous suspensions of each of the powder preparations described under (i) were introduced into the reaction chamber of the Van Slyke microapparatus and deaminated in the usual way The preparations did not show any significant difference in amino-N content, the mean value of which was 0 15 % (range 0 13 to 0 17 %) for the dry materials

(iii) *Titration data* (*a*) *Silk fibre* Samples (2 0 g) of degummed silk (rendered ash-free by soaking in 0 02N-acetic acid for 24 hr and subsequent washing with distilled water) containing 8 9 % of moisture, were allowed to reach equilibrium at 20° C with 100 ml of HCl or NaOH of various concentrations After 48 hr the amounts of acid or alkali left in the solutions were determined by means of the glass electrode for pH values between 2 and 12, and by titration for values outside this range. The amount taken up by the silk was then found by difference (*b*) *Fibroin powder* Fibroin powder was prepared as described under (i), except that the powder after precipitation was washed with dilute HCl in order to remove traces of residual copper (shown by testing with Na diethyldithiocarbamate) and then exhaustively washed with distilled water. Titration data were obtained by the same procedure as employed for the fibrous material.

The results, shown graphically in figure 2, indicate that the acid- and alkali-combining power of fibroin powder is slightly less than that of the initial fibrous material, it may be taken that the degradation occurring during regeneration is negligible The data show that the maximum amount of HCl bound is about 15 ml of N-solution per 100 g of fibroin. This figure appears reasonable since the free amino-N, 0 15 %, is equivalent to 11 0 ml of N-acid and the basic non-amino-N would somewhat increase this figure The results are in agreement with those of Harris (1941), the higher value given by the earlier work of Lloyd & Bidder (1935) was probably due to extraneous matter present in the fibroin used. According to the work of Cohn and his associates (1934, 1938) on zein, ovalbumin, globin, and other proteins, the lack of combination with acid or alkali between pH 3 3 and 8 0 is due to the almost complete absence of free carboxyl groups, and the flatness of

the curve between pH 7 and 9 to the small number of glyoxaline groups. Above pH 8, combination with alkali is attributed to the tyrosine groups, this confirms the results of Harris (1941) From the tyrosine content of fibroin (10 6 %), a maximum combination of 59 0 ml. of N-NaOH is to be expected The curve shows that this degree of combination is apparently reached at about pH 12 9, if the curve had been taken above pH 13, an inflexion might have been expected to occur between tyrosine combination and combination due to —CO—NH— linkages The lack of acidic and basic groups in fibroin obviously precludes a well-defined isoelectric point

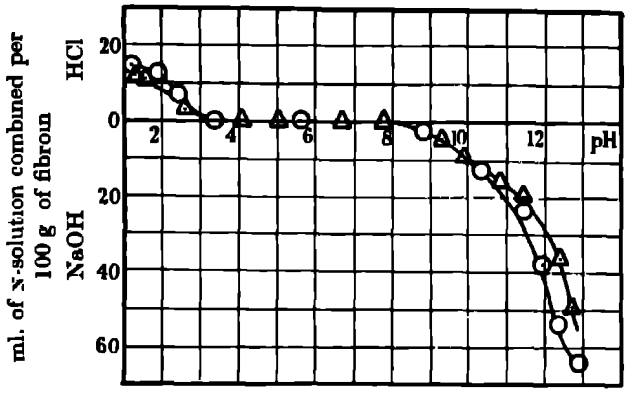


FIGURE 2 Titration curves of fibroin powder (△) and silk fibre (○).

An interesting feature of the solutions of fibroin prepared as described above was their behaviour on acidification. For the solutions prepared with a period of 3 min between dissolution in 6/8 solvent and neutralization, about two-thirds of the protein material was precipitated as a mucilaginous mass on adjustment of the reaction to below pH 5 This precipitation resulted in a clarification of the solution, the residual protein solution being water-clear It appeared, therefore, that there existed in the initial solution two fractions, one of protein material in a dispersed state and the other of protein material in a water-soluble form When the solution of fibroin in 6/8 CuEn is left for various times before the neutralization and subsequent dialysis procedure, the proportion of protein not precipitated by acidification to pH 3 increases during the first 10 min and then remains constant This effect of the time factor is shown in table 3, for which 5 0 g. of air-dried fibroin was dissolved in 50 ml. of 6/8 solvent and the solution left for various times before neutralization The highest proportion of water-soluble (as distinct from dispersed) protein produced under the conditions used appeared to be about 50 % and it was of immediate interest to ascertain whether or not this yield could be increased. The fibroin is primarily maintained in solution by the CuEn reagent and, as is shown below, molecular weight determinations indicate that complete molecular dispersion occurs within 3 min. under these conditions, i e the reactants of the system are available probably immediately after dissolution of the fibroin. The

data given in table 3 point to the establishment of a half-way equilibrium between dispersed and water-soluble forms of fibroin (as CuEn complexes) and this equilibrium should be shifted by a change in the concentration of CuEn reagent This actually occurred, for when fibroin (5 0 g) was kept for 10 min in a 12/16 solvent (50 ml) before neutralization, the resulting protein solution gave no precipitate when acidified to pH 3

TABLE 3 EFFECT ON SOLUBILITY OF VARIATION IN TIME BETWEEN DISSOLUTION AND NEUTRALIZATION FOR 5 G OF FIBROIN IN 50 ML. OF 6/8 SOLVENT

time between dissolution and neutralization (min)	fluidity (reciprocal poises)	% soluble at pH 3
1	2 5	15 0
3	3 0	33 9
4	3 2	37 3
10	5 2	49 6
15	6 4	51 0
30	7·5	53 0
60	9 2	51·0

The water-soluble protein in 2 to 3 % concentration is precipitated completely by (i) an equal volume of alcohol or acetone, (ii) the addition of 2 % of trichloroacetic acid, or (iii) saturation with inorganic salts, e g NaCl, $(NH_4)_2SO_4$, $MgSO_4$, $NaPO_3$ However precipitated from its aqueous solution, the protein becomes insoluble in water Since most proteins that are precipitated by acetone at temperatures below 5° C will dissolve again in water, precipitation of the water-soluble fibroin was carried out at −60° C, the precipitate was again completely insoluble in water Examination of solutions of 50 % of dispersed and 50 % of water-soluble ('50/50 mixture') or of 100 % water-soluble protein gave the following results

(*a*) *Stability* Neutral solutions 1 to 3 % in total protein and containing approximately equal parts of dispersed and water-soluble forms of fibroin (50/50 mixture) are stable at room temperatures for about 14 days after preparation and then set to a solid gel of insoluble protein The water-soluble form of fibroin in 1 to 2 % solution at room temperature is less stable than is the 50/50 mixture and gels within about 4 days of preparation, or less if the solution is maintained at an acid reaction A remarkable property of the water-soluble preparation is that heating to 100° C for 10 min does not result in any coagulation of the protein or in the formation of an opalescence in the liquid. The 50/50 mixture also shows no early sign of coagulation on heating to 100° C

(*b*) *Molecular weight* * The insolubility of fibroin in the usual protein solvents has so far prevented the direct determination of its molecular weight Two values are reported in the literature, viz 300 by Herzog (1928) and 217,700 by Bergmann

* The term 'molecular weight' implies *mean* molecular weight and is used in this sense throughout the paper

& Niemann (1938), the former value was obtained from data for depression of the melting-point of resorcinol by fibroin (which is considerably degraded by the dissolution) and the latter by a calculation based on a histidine content of 0·07 % determined by Vickery & Block (1931). The present preparations of fibroin are well adapted to osmotic-pressure measurements since there is no combination with acid or alkali. Solutions at pH 7 and free from electrolytes were examined by the static method of Burk & Greenberg (1930), the apparatus being immersed in a thermostat at 20° C and the manometric level observed at intervals by a cathetometer. Equilibrium was reached in 5 to 6 days. A check on the method was provided by an examination of edestin hydrochloride (freshly prepared from hemp seed) in water. The osmotic pressure, recorded in centimetres of water, for the edestin solution (pH 3·63) agreed with that derived by extrapolation from the curve given by Burk & Greenberg (1930) for the osmotic pressure of edestin in 6M-urea solution at various values of pH. The osmotic pressures of preparations of the 50/50 mixture (0·5 to 3·0 % in total protein) were found to be independent of type of membrane (cellophane or parchment) and of the direction of approach to the equilibrium hydrostatic pressure (Adair 1925, Roche, Roche, Adair & Adair 1932). The results (figure 3) show that for concentrations below 1·5 % a linear relationship is followed (the deviation for concentrations of 1·5 to 3·0 % is slight) and hence for such solutions the Van't Hoff relationship, M w. $= RTc/p$, is applicable. These results were obtained on fibroin that had been dissolved in the 6/8 solvent for 10 min. before neutralizing and it is possible that a small amount of degradation of the protein had occurred. Accordingly, measurements were also made on fibroin that had been dissolved for 3 and 5 min., respectively, the neutralized solutions had fluidities of 5·0 and 7·5 reciprocal poises. As is seen from figure 4 (points indicated by arrows), the molecular weights determined on these two solutions are slightly and progressively higher than is that obtaining for the 50/50 mixture (10 min. dissolution). The question of immediate interest is the comparison of the molecular weights of the dispersed and soluble protein. After separating the dispersed fibroin from a 50/50 mixture by precipitation with dilute acid measurements were carried out on solutions of the soluble product alone. A value of 33,000 was obtained and this is not significantly different from that for the dispersed protein, a 50/50 mixture giving the same value ± 1000 (figures 3, 4). The complete range of molecular weights (33,000→7,000) and fluidities (5·0→44·0) were determined for fibroin and its degradation products (alkali hydrolysis) by the usual 10 min. dissolution procedure and are shown in figure 4. The relationship between molecular weight and fluidity is linear, as is also the amino-N–fluidity relation (figure 5) over the first half of the fluidity range and for both acid- and alkali-hydrolyzed fibroin. Above fluidity 44, silk has lost its fibrous properties, i.e. it disintegrates to a powder, at this point the molecular weight is about 7,000. This is in agreement with the work of Carothers (1940) who showed that when a polymeric series is ascended, fibrous characteristics first appear when the average chain-length corresponds to a molecular weight of about 10,000.

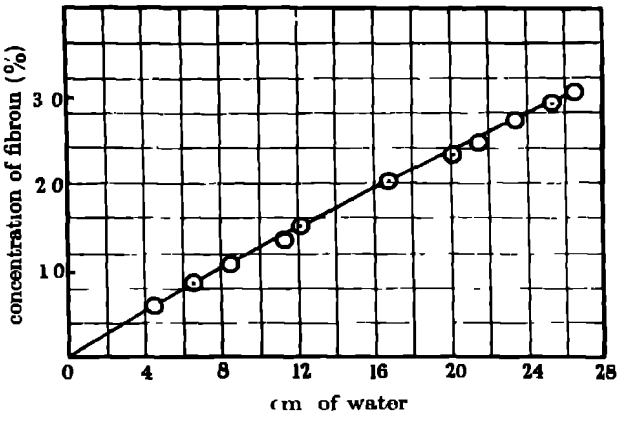


FIGURE 3 Relation between osmotic pressure (in cm of water) and concentration of solutions of 50 %-renatured fibroin

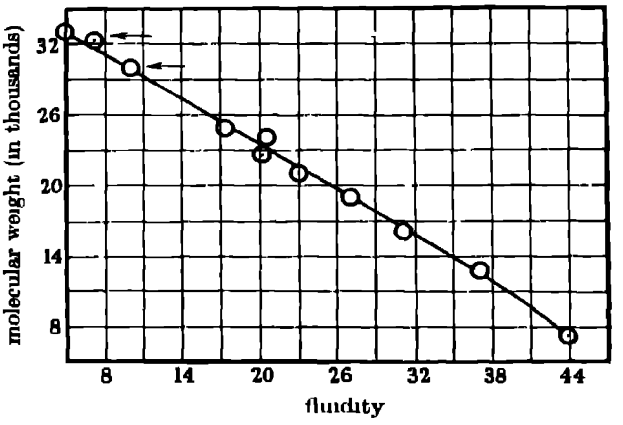


FIGURE 4 Relation between fluidity (in reciprocal poises) and molecular weight of fibroin and degradation products formed by alkaline hydrolysis

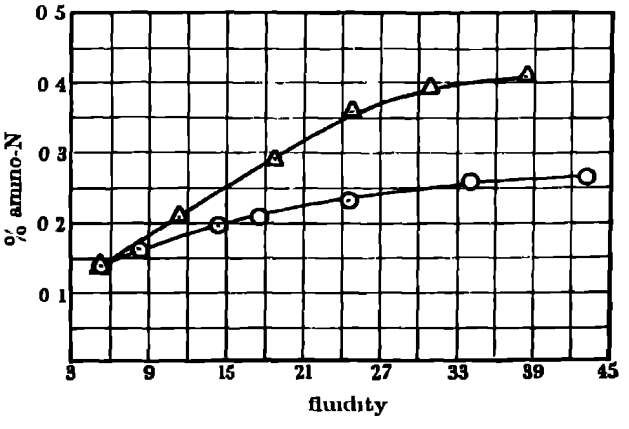


FIGURE 5 Relation between fluidity (in reciprocal poises) and amino-N content of fibroin and degradation products formed by alkaline (⊙) and acid (△) hydrolysis.

(c) *Titration curves: ultra-violet absorption spectra: tyrosine-hydroxyl groups.* A 50/50-mixture preparation of fibroin was titrated with 0·1N-NaOH by means of the glass electrode. The solution could not be titrated on the acid side since the dispersed protein separated out at pH 4 to 5. For the titration of the soluble fibroin, the mixture was first acidified and the precipitate removed in the centrifuge. The solution was then adjusted to pH 7 and one aliquot was titrated with 0·1N-NaOH and another with 0·1N-HCl. All the titrations were performed four times and the data from one of these experiments are shown in figure 6. Comparison of the two curves is possible only above pH 5 for which it is seen that the dispersed fibroin combines with more alkali over the range pH 7 to 11·5 than does the soluble. Above pH 11·5, the uptake of alkali is about the same. The explanation of this divergence appears at first sight to be related to the observations of Crammer & Neuberger (1943), who determined the ultra-violet absorption spectra of tyrosine and tyrosine-containing proteins. These workers found that native ovalbumin contained none or few ionizable phenolic groups and they postulated that in the native protein the phenolic groups are bound in a linkage, possibly a hydrogen bond. On denaturation, this bond is broken and the phenolic groups ionize to a greater extent. We have found, however, that this difference in ionization does not obtain between dispersed and soluble fibroin. Solutions of soluble fibroin and of 50/50 mixture, both 0·1 % in total protein, were examined for ultra-violet absorption by means of a sector spectrophotometer with a hydrogen discharge tube as source of ultra-violet light. The curves obtained for molecular extinction coefficient against wave-length with the solutions at pH 11·0 and 12·0 (by glass electrode) are shown in figure 7, from which it is seen that there is no significant difference in ultra-violet absorption between the two forms of fibroin over the range of 260 to 310 mμ. That this absence of change in the availability of tyrosine is not incompatible with the titration data is evident from the following considerations. The number of phenolic groups ionized per molecule of fibroin at pH 11 or 12 can be calculated from the absorption curves for tyrosine given by Crammer & Neuberger. At pH 11, the molecular extinction coefficient $\epsilon \times 10^{-4}$ is 3·60 at $\lambda = 295$ mμ, at which wave-length the value $\epsilon \times 10^{-4}$ for tyrosine at pH 11 is 0·23. Hence the number of phenolic groups ionized per molecule is $3{\cdot}60/0{\cdot}23 = 15{\cdot}6$. The corresponding value for $\lambda = 300$ mμ is $2{\cdot}90/0{\cdot}197 = 14{\cdot}7$. Since there are about 20 tyrosine residues per molecule, it follows that about 75 % of the phenolic groups are ionized at pH 11. The titration data, however, indicate that at pH 11 soluble fibroin combines with 60 ml. of N-NaOH per 100 g. and dispersed fibroin with 86 ml. Since the whole of the tyrosine if present in titratable form would combine with 59 ml. of N-NaOH, it follows that at pH 11 about 15 ml. of N-NaOH are combined with other acidic groups in the soluble protein and about 40 ml. in the dispersed protein. Thus there appears to be a difference in the base-combining power of the two forms of fibroin that is due to differences in availability of some group or groups other than phenolic, the nature of which is at present not clear. The equal availability of tyrosine phenolic groups in the two forms was confirmed by an examination of

their rates of oxidation by potassium ferricyanide solution (cf. Mirsky & Anson 1936).

The rate of oxidation of tyrosine under standard conditions was first investigated 0·5 ml of 0·1 % tyrosine + 1 0 ml. of N/20-carbonate-bicarbonate buffer at pH 11 + 0 2 ml of M/2-potassium ferricyanide were left for set times. The amount of ferrocyanide produced was then determined by adding 5 ml. of the ferric iron

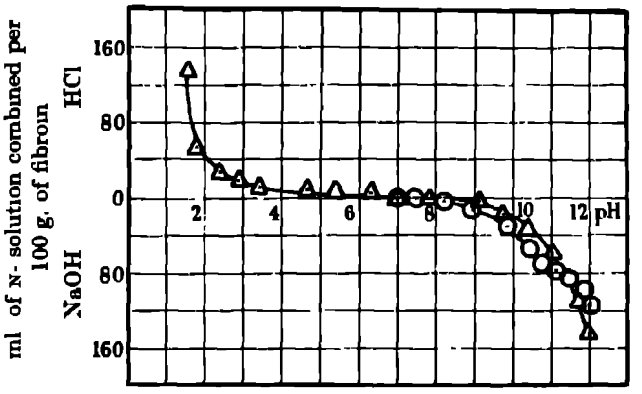


FIGURE 6 Titration curves of 100 %- (△) and 50 %-renatured (○) aqueous fibroin solutions

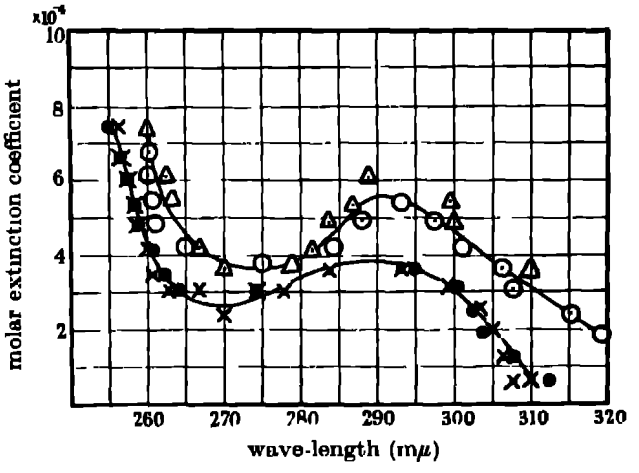


FIGURE 7 Absorption spectra of renatured fibroin at pH 11 (●) and pH 12 (∧) and of 50 %-renatured fibroin at pH 11 (×) and pH 12 (○)

solution of Folin & Malmros (1929) and comparing the blue colour with that given by a standard prepared by adding 1·0 ml. of 0·1 % potassium ferrocyanide to 5 ml of ferric iron solution and 0 7 ml of water The curve of ferrocyanide produced against time showed two flat portions The tyrosine was oxidized rapidly during the first 20 min and thereafter the amount of ferrocyanide produced remained constant (at 4 mg) for 2 hr , corresponding to 2 atoms of oxygen per molecule of tyrosine. The second flat portion occurred after 24 hr. at 6·6 mg of ferrocyanide.

Whilst intact silk filaments had no action on the potassium ferricyanide solution, both the soluble and the 50 %-dispersed fibroin showed a marked reducing action The extent of reduction was the same for both types of protein and was about twice as great as obtained for the equivalent amount of tyrosine under the same conditions. Thus with a reaction time of 1 hr , the tyrosine in the fibroin molecule apparently utilized about 4 atoms of oxygen per residue There is a possibility that some amino-acid residue other than tyrosine is responsible for the extra reducing action, but it was found that serine, phenylalanine, leucylglycylglycine, proline, and hydroxyproline did not reduce ferricyanide under the above conditions

(*d*) *Optical rotation* Preparations of the soluble fibroin and of the 50/50 mixture were examined in a polarimeter in a 20 cm tube, jacketed with water at 15° C, a sodium vapour lamp being the source of light The results are given in table 4.

TABLE 4 OPTICAL ROTATION OF '50/50 MIXTURE' AND OF COMPLETELY WATER-SOLUBLE PREPARATIONS OF FIBROIN

solution		pH	conen of fibroin %	rotation	$[\alpha]_D^{15}$
50/50 mixture	prep A	7 0	3 00	−3 39'	−56 4°
	prep B	7 0	3 03	−3 35°	−55 3°
	prep C	7 0	1 51	−1 70"	−56 1°
	prep D	7 0	2 95	−3 31°	−56 2°
	prep E	7 0	3 24	−3 63"	−56 0°
	prep F	7 0	2 06	−2 31°	−56 1°
soluble fibroin	prep G	7 0	1 78	−1 90°	−53 4°
	prep H	7 0	1 14	−1 20'	−52 8°
	prep I	7 0	1 08	−1·98°	−53 5°
	prep J	5 0, 6 1, 7 0, 9 2, 10 0, 11·1	1 20	−1 27'	−53 0'
	prep. K	7 0	1 05	−1 11°	−53 2'
preparation K after heating to 100° for 10 min				−1 12°	−53 3°

No determinations were made below pH 5 0, owing to the development of an opacity in the solution at low pH values The fact that the rotation of solutions of soluble fibroin remains unchanged after the solution has been heated to 100° C and then cooled emphasizes its remarkable stability to heat. There is no sign of coagulation, the solution remaining quite clear

(*e*) *Primary effect of tryptic hydrolysis* The work of Lin, Wu & Chen (1928) on ovalbumin and of Linderstrøm-Lang (1939) on lactoglobulin indicated that the enzymic digestion of proteins proceeds more rapidly with the denatured than with the native form. This was confirmed by Bernheim, Neurath & Erickson (1942) with horse serum-albumin and -globulin In the present work on fibroin, the tryptic digestion of a 50/50 mixture was compared with that of a solution of soluble fibroin of the same concentration of total protein With systems of 100 ml of a 1 % solution of protein in M/30-phosphate buffer at pH 7·8 and containing 25 mg of trypsin (B P), no significant difference in the rate of hydrolysis at 37° C (measured

by liberation of amino-N) was observed between the two forms of protein. After digestion for a few minutes, however, both solutions deposited a white, amorphous material and this deposition continued at a rate that was greater with the 100 %-soluble than with the 50/50 mixture The precipitated material (production of which ceased after digestion for about 2 hr.) was insoluble in dilute acids and alkalis and had a remarkably low content of tyrosine It could be made water-soluble by the CuEn regeneration method but the product was invariably far less stable than were regenerated preparations of fibroin Some of the preparations of the regenerated precipitate were sufficiently stable to permit examination of their molecular complexity by fluidity determinations, but in order to measure their osmotic pressure it was necessary to add to the precipitated material a definite amount of fibroin of known molecular weight before regeneration and, further, to make the dialysed solution 6M in urea in order to give a solution of adequate stability The material deposited at the end of the 2 hr. digestion period had a molecular weight of the order of 2000 to 3000, whilst that deposited during the first 10 min. gave a value of 8000 to 10,000 Preparations of the latter type, which had been freed from trypsin by washing in 0·1N-HCl at 0° C before examination for molecular weight, etc , contained 1 % or less of tyrosine and no proline

(*f*) *Viscosity* Aqueous solutions of the 50/50 mixture and of soluble fibroin were prepared by the usual methods and the concentrations of the solutions were adjusted so that both types were 0·76 % in total protein. The viscosity of these solutions was measured in a viscometer similar to that used by Garrett & Howitt (1941) excepting that the capillary diameter was approximately one-quarter of that of the viscometer described. The viscosities obtained were 1·46 and 1·75 centipoises for the solutions of soluble fibroin and 50/50 mixture, respectively

(*g*) *Films and filaments* As stated above, it was not found possible to precipitate the soluble fibroin in a water-soluble form by the use of protein precipitants However, a thin layer of the aqueous solution on a mercury surface in a vacuum desiccator over phosphorus pentoxide yielded a transparent film of fibroin that was soluble in water. When slightly moistened, the film was fragile but could be stretched to about three times its original length. This stretching resulted in a large increase in tensile strength and resistance to tear, the stretched film being completely insoluble in water, and strongly birefringent when examined through crossed Nicols Both stretched and unstretched films gave an X-ray picture identical with that of fibroin powder and corresponding to the spacing shown by degummed silk fibres, i.e. the β-keratin pattern of Astbury (it was established that the process of examining the unstretched, water-soluble film by X-rays did not produce insolubilization of the film) In order to obtain water-soluble films, it is necessary to centrifuge the solution before use and to carry out the vacuum drying as rapidly as possible: failure to comply with these conditions results in the production of slightly opaque films, which are not, or only slightly, soluble in water, do not readily stretch and, when stretched, do not show birefringence or any apparent increase in tensile strength.

Concentrated (about 10 %) solutions of fibroin could be prepared by partial drying in a vacuum and from such solutions fine fibrils could be drawn that showed a tensile strength approximately one-third to one-half of that of natural silk filaments. X-ray examination of these fibrils indicated that the low tensile strength was due to lack of molecular orientation

(*h*) *Crystallizability.* Some eighty or more attempts were made to crystallize fibrinogen but, despite numerous variations in experimental technique, no unqualified success was achieved The nearest approach to a crystalline product, obtained by slow separation from a glycerophosphate buffer at pH 6 in presence of zinc ions, is shown in figure 12, plate 5; this product was only partly water-soluble.

3 Discussion

A *Reasons for believing the water-soluble preparation to be renatured fibroin*

During the past decade, much attention has been directed to the conversion of globular proteins of ready solubility in water or dilute solutions of electrolytes into fibrous proteins that are characterized by their relative insolubility. Such a conversion has been achieved with certain proteins (e g. ovalbumin, casein, peanut globulin) in which the formation of a fibre is entirely divorced from the natural function of the proteins. There are, however, many reasons for considering that it would be more fruitful, at least from a strictly academic viewpoint, to consider instances where the conversion of a water-soluble into a fibrous protein represents the fundamental role of the protein. Such a protein is fibroin, which occurs in a water-soluble form in the silk gland and, during the process of extrusion through the spinneret, is converted into an insoluble fibrous form in order to construct the habitat for the next metamorphous stage of the silkworm Although the possibility of an enzymic mechanism for this conversion has not been disproved (see Gilson 1904), most workers have concluded that the insolubilization is due to the mechanical stresses set up on extruding a viscous mass through a fine orifice (see, for example, Foà 1912). The truth of the latter explanation would imply that the insolubilization of the silk-gland protein is a relatively simple process in which linkings such as those of primary valency are not involved: the possibility of its reversal in the laboratory is therefore not to be excluded. Treatment of the insoluble fibroin with cupri-ethylenediamine reagent converts the protein partly or wholly into a water-soluble form; with relatively low concentrations of the reagent, part of the fibroin is produced in a dispersed state that is readily coagulated by acidification The data obtained indicate that the reaction by which the CuEn-fibroin complex changes to the CuEn complex of the soluble form follows the normal chemical laws. Regeneration of the cupri-ethylenediamine complex by neutralization yields unchanged fibroin in a dispersed state and changed fibroin in a water-soluble form. The main point of interest is the nature of this conversion. It has been established to a reasonable extent that the conversion is attended by no significant change in amino-N content or titration data and that no liberation of

diffusible amino-acids or peptides occurs The change in molecular weight is very small and, from the aspect of long-chain molecules of a more or less repetitive pattern, the difference between the dispersed and water-soluble forms is negligible. It appears, therefore, that we are concerned with a rearrangement of the intact molecule, which, according to the equal accessibility of tyrosine-hydroxyl groups in dispersed and water-soluble forms, must be of a relatively simple nature. The values of the viscosity of the 50/50 mixture and of the water-soluble form indicate that solubilization is attended by a diminution in the degree of dyssymmetry of the molecule and the picture that can be tentatively drawn is one of a folding of a molecule that is accepted to be in the form of a long extended chain. Such a conclusion leads to the following interpretation of the general mechanism of the reactions attending dissolution of the fibrous protein in cupri-ethylenediamine reagent

The natural silk filament consists of extended polypeptide chains in high degree of alinement along the axis of the fibre, the nature of the chain being such that intermolecular binding, an important factor in fibre strength, is mainly due to hydrogen bonds between adjacent CO and NH groups of the peptide linkings On treatment with cupri-ethylenediamine, these bonds are broken by complex formation between CuEn and .NH groups. Subsequent neutralization with acetic acid as soon as possible after dissolution results in the decomposition of the complex and hence in the liberation of the long-chain molecules as individual units. These units (which form the dispersed stage) are relatively unstable and readily aggregated by, for example, dilute acids. If, however, the complex is not neutralized until some minutes after formation, it is structurally modified to a folded, water-soluble form to an extent dependent on time and concentration of cupri-ethylenediamine The folded chain has its folds internally cross-linked by hydrogen bonds between pairs of ·NH and .CO groups (cf. Buswell, Krebs & Rodebush 1940, Alexander & Rideal 1941), so that external intermolecular linking does not readily occur, hence the folded form is less sensitive than is the unfolded form to the effect of external influence such as hydrogen-ion concentration.

The reverse process of insolubilization is effectively demonstrated by the films of water-soluble fibroin. Stretching of a strip of moist film results in an extension to about three times its original length and the simultaneous acquisition of tensile strength and insolubility It appears that this mechanical stretching results in the unfolding of the folded molecules, which then aggregate in a manner similar to that occurring in the natural silk filament; indeed, it is difficult to avoid the conclusion that the stretching and insolubilization of the moist strip must be akin to what is happening when the silkworm spins its fine filament

An obvious corollary to the above conclusions is that the water-soluble fibroin is similar to or identical with the water-soluble protein that is present in the silk gland and yields the natural silk filament Soluble fibroin taken directly from the gland is probably invariably contaminated with sericin (the two proteins exist as separate sols in the gland), but descriptions of such preparations given by Meyer & Jeannerat (1939), Foà (1912), Ongaro (1933), and others strongly support the

identity of the regenerated fibroin with the natural, water-soluble product. It is therefore probable that the cupri-ethylenediamine treatment affords a reversal of the insolubilization process that occurs during the spinning of fibroin filaments by the silkworm

The above reversible solubilization of fibroin has an important bearing on the phenomenon of denaturation According to Neurath, Cooper & Erickson (1942), the main changes accompanying denaturation are (*a*) a loss of biological activity, (*b*) a decrease in solubility, (*c*) a loss of crystallizability, (*d*) an apparent increase in the number of specific groups, e g thiol, dithio, or phenolic groups, and (*e*) an increase in the relative viscosity of the solution of a protein and a corresponding decrease in its diffusion constant, indicative of an increase in apparent molecular asymmetry. It has been shown that insolubilization of fibroin is attended by an increase in molecular asymmetry. Fibroin has no specific biological activity (in agreement with its function as a building material), whilst the lack of change in available tyrosine groups is probably related to its simplicity of structure (see below). With regard to loss of crystallizability, it has been shown that the water-soluble form can be separated in a form that possesses the near-crystal habits of native globular proteins when in an impure state. The above criteria thus afford some justification for regarding the insolubilization of water-soluble fibroin as a process of denaturation. Moreover, it has been shown by various workers that the optical rotation of proteins increases on denaturation. Thus Young (1922) found that on denaturation ovalbumin increased in specific rotation from $-36{\cdot}6°$ to $-40{\cdot}4°$ and serum-albumin from $-78\ 6°$ to $-89{\cdot}6°$. With the aqueous preparations of fibroin, it was noted that the optical rotatory power is independent of pH and of protein concentration, in agreement with the observations of Barker (1933) on albumin Barker considered that 'the optical rotatory power is the only property which at present is known to be suitable for quantitatively characterizing a denatured protein' It can be calculated from the values obtained for the 50/50 mixture and the water-soluble preparation that the water-soluble form has $[\alpha]_D^{15} - 53\ 1°$ and the dispersed form $[\alpha]_D^{15} - 58\ 9°$ There appears to be justification, therefore, in regarding the dispersed form of fibroin (probably identical with the natural silk filament) as denatured fibroin and the water-soluble form (probably identical with the fibroin as it exists in the silk gland) as the renatured or native protein.

The films of water-soluble fibroin give an X-ray diagram of the β-keratin type, which is not changed by the process of insolubilization induced by stretching It follows that the α-β-keratin transformation described by Astbury is not involved in the process of denaturation of fibroin (see Bailey, Astbury & Rudall (1943) and Neurath *et al.* (1944, p. 234)) Note also that, according to Spiegel-Adolf & Henny (1941), ultra-violet irradiation of serum-albumin causes denaturation without any significant changes in the X-ray picture.

Fibroin, owing to its relative simplicity of structure, appears to represent the ideal example of denaturation. In fact it can be claimed that denaturation is the

main biological function of this silk protein and lends emphasis to the necessity for studying denaturation as a natural phenomenon rather than as an artificial process such as the heat coagulation of ovalbumin. In this connexion, the conversion of blood fibrinogen to fibrin appears to be essentially an enzymically controlled process of denaturation and it would be reasonable to designate the soluble protein of the silk gland that is the precursor of the filament protein as fibroinogen, in analogy with fibrinogen.

B. *A possible mechanism for the renaturation of fibroin based on a special conception of its structure*

The amino-acid composition of fibroin (see table 2) and the general spatial arrangement of the molecule as given by the X-ray investigations of Astbury (1934) and others indicate that the protein molecule, as present in the silk fibre, is composed of an extended long chain consisting mainly (over 70 %) of glycine, alanine, serine, and tyrosine residues. The molecular weight of about 33,000 to 34,000 and the mean residue weight of 78 mean that there are some 420 to 430 residues in the molecule, of these, the less frequent residues are represented by histidine (1), lysine (1), arginine (2), leucine (2), threonine (4) and proline (4) The reactive groups of the molecule are mainly tyrosine (phenolic) hydroxyl and serine (alcoholic) hydroxyl, the number of free amino- and carboxyl groups being very small. This representation of the fibroin molecule is appropriate for the construction of the silk filament but affords no information in an understanding of the existence of a water-soluble form. The solubilization of fibroin has been correlated above with a folding of the molecule and, if it is accepted that both the insoluble and soluble forms consist of essentially the same polypeptide chain with a definite arrangement of amino-acid residues, it follows that this arrangement of amino-acid residues must be such that a folding may occur to produce one definite spatial arrangement of the chain Up to the present, the only theory regarding the distribution of the amino-acid residues along the length of the polypeptide chain is that of Bergmann & Niemann (1938) According to this theory, (*a*) the total number of any particular residue or of all the residues in the protein molecule may be expressed as $2^n.3^m$, where n and m are whole numbers, and (*b*) the residues are spaced at regular intervals along the molecular chain. With regard to (*a*) the evidence available does not afford unanimous support and it is necessary to accept the hypothesis that there is some repetition of pattern in the protein molecule. With regard to (*b*), it is reasonable to accept the premise that each residue plays an individual or composite part in the physico-chemical and biological properties of the protein molecule of which it is a constituent; in other words, the arrangement of residues in the polypeptide chain is a rigidly defined characteristic of a protein. This view is not in harmony with a theory of periodic distribution of the residues, which greatly restricts the occurrence of any specific function of a residue that is dependent on the particular location of that residue in the chain Thus, according to the Bergmann-Niemann theory, the twenty tyrosine and four proline residues

are regularly spaced along the fibroin polypeptide chain, but this is completely negatived by the fact that limited tryptic digestion splits the molecule into parts the average chain-length of which is nearly one-third of that of the parent molecule and which contain no proline and only about 1 % of tyrosine. The inference is that all or most of the tyrosine residues and all four of the proline residues are situated in short parts of the chain at one-third and two-thirds along the length of the chain from one end. This fact, together with the isolation of serylprolyltyrosylproline from fibroin by Abderhalden & Bahn (1932), point to a special significance of proline in the folding of the chain Stereochemical considerations show that, where a proline residue is present, the chain must be turned through an angle of about 90° and that a prolyltyrosylprolyl grouping can produce a complete fold in the chain. If it is assumed that each of the two tyrosine-rich parts of the chain contain two proline residues, it follows that the long fibroin chain may fold at these two molecular hinges to give a three-membered grid.

The above theory leads to a tentative explanation of the renaturation of fibroin by cupri-ethylenediamine. In the vicinity of the proline residues, the structure of the fibroin-cupri-ethylenediamine complex may be represented as shown in figure 8 Since rotation round alternate —CO—CH*R*— bonds has to be assumed along the whole chain in order to explain the formation of the complex, it is reasonable to postulate rotations about the —CO—CH*R*— bond indicated in figure 8. On constructing a solid model of this part of the chain, it was found, however, that the two CuEn groups could not be accommodated in this structure. It was considered that the molecule could best be stabilized in the folded form by assuming that the remaining En is displaced from one of the CuEn radicles, as shown in figure 9. A solid model of this structure indicated that the four donating nitrogen atoms were all in one plane (the co-ordination number of copper is six, and the remaining two positions are occupied by water molecules as with cupri-ethylenediamine itself) and that a copper atom just fitted into the assigned position. The resemblance of the structure to that of haemoglobin, chlorophyll, monastral blue, etc., is noteworthy. The theory now suggested for the solubilization (or renaturation) and insolubilization (denaturation) of fibroin is that the native or regenerated protein exists as a 3-limbed polypeptide chain that spontaneously (keeping of the solution) or mechanically (stretching of the moist film) changes into the extended chain. The extended-chain form (referred to above as the dispersed form) initially consists of single molecules that aggregate to give insoluble material at a rate dependent on the environmental conditions This introduces a further aspect of denaturation that has been the subject of much controversy, viz , the justification for the separation of the processes of denaturation and the subsequent flocculation or coagulation. The above experiments on fibroin support the separation of the two processes, i.e. the denaturation of fibroin refers solely to the unfolding of the folded molecule The formation of molecular aggregates is a common phenomenon with macro-molecules and there is no reason why protein molecules should not behave as typical members of this class of substances.

Disaggregation results in a dispersal or solubilization of the insoluble aggregate; an example is the production of water-soluble cellulose by Lieser (1940). Water-soluble fibroin has at least a tendency to aggregate in a manner similar to the crystallization process of the globular proteins (figure 12, plate 5). A general scheme of these changes may be represented by.

molecular aggregate $\rightleftharpoons$ single molecules of native protein $\rightleftharpoons$ unfolded or denatured protein molecule $\rightleftharpoons$ aggregate of unfolded protein molecule

and, bearing in mind the data that have been obtained for the molecular weights of various proteins under different conditions, it would appear that this scheme is of general applicability to proteins.

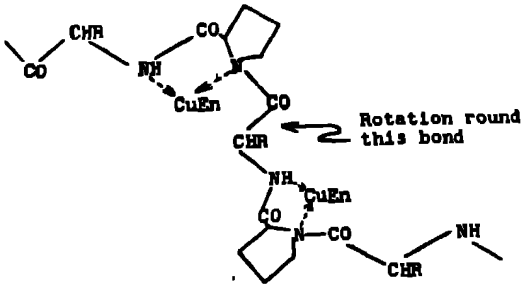


FIGURE 8. Mechanism of chain folding in renaturation of fibroin

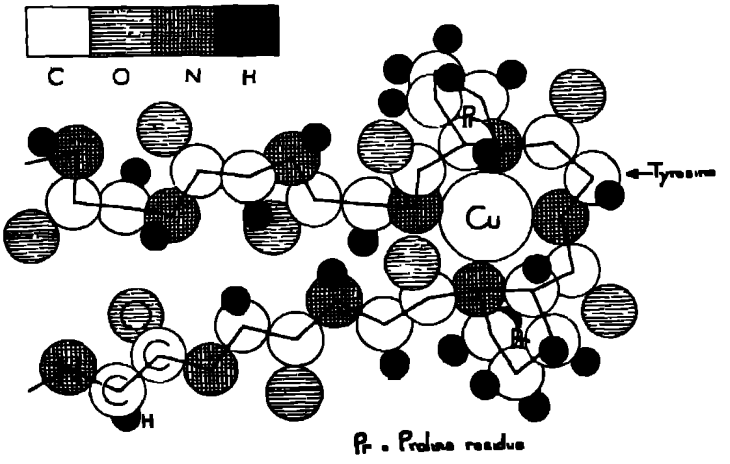


FIGURE 9 Structure of the 'molecular hinge' during the renaturation of fibroin by cupri-ethylenediamine

The stability of the water-soluble form of fibroin when its solutions are heated to 100° C is noteworthy; the complete absence of conversion into the denatured form when the protein is heated to 100° C for 10 min. is shown by the constancy of the optical rotation. The marked difference between fibroin and proteins such as ovalbumin is probably due to the comparative simplicity of the structure of

fibroin and it is of interest to inquire in a preliminary manner how far the mechanism of denaturation and renaturation of fibroin is applicable to other proteins.

Chibnall, Astbury and others have concluded that globular proteins are probably multilaminar in structure, each lamina being formed by the folding backwards and forwards of a single polypeptide chain. The chains that fall into a plane (the possible number of which is equal to the number of limbs in the lamina) at right angles to the lamina form a 'polypeptide grid' (Astbury), which contains linkings such as dithio. The experimental evidence points to a unilaminar structure for fibroin, the unfolding of which yields a single polypeptide chain, the ideal structure for fibre formation The assumption of an unfolding mechanism for the multilaminar proteins, however, implies that a molecular sheet or ribbon is formed, the individual polypeptide chains of which are linked by, for example, dithio bonds True fibre formation would be possible by prior or subsequent rupture of the dithio and similar linkings and, with respect to blood-fibrinogen, interest in such a possibility is enhanced by consideration of the recent conclusion of Lyons (1945) that the thrombase system contains a component capable of doing this, together with the earlier observations of Barrett (1920) on the filamentous nature of fibrin. A final point of interest concerns the solubility of the degradation products afforded by enzyme- or acid-hydrolysis of fibroin, a question that at present is being further investigated These products, even when as low as 20-peptide, are insoluble in water or dilute acids or alkalis. 'Regeneration' by the cupri-ethylenediamine method, however, gives water-soluble products that spontaneously revert rapidly to the insoluble form. This process of regeneration cannot involve a folding mechanism but is due merely to a separation of aggregated polypeptide chains, the resulting solution exhibiting what appeared to be 'coagulation' unaccompanied by prior denaturation In addition, the experimental work on fibroin has indicated that slight degradation of the fibroin molecule, i e. that attended by a loss of a few (probably terminal) amino-acid residues, does not prevent denaturation or renaturation. This implies that, at least for fibroin, the process of denaturation is not necessarily restricted to the completely intact or native protein.

Grateful acknowledgement is made to Dr R J. B. Marsden and Dr F. Howlett of the British Cotton Industry Research Association for stimulating discussions concerning macro-molecules, and to Miss M. Evans and Mr W. Wainwright for many careful analyses and technical assistance

References

Abderhalden, E. & Bahn, A. 1932 *Hoppe-Seyl. Z.* **210**, 246.
Adair, G. D. 1925 *Proc. Roy. Soc.* A, **108**, 627; **109**, 292.
Alexander, A. E. & Rideal, E K. 1941 *Nature*, **147**, 541.
Astbury, W. T 1934 *Cold Spr. Harb. Symp Quant Biol.* **2**, 15
Bailey, K., Astbury, W T & Rudall, K. M. 1943 *Nature*, **151**, 716.
Barker, H A. 1933 *J. Biol Chem.* **103**, 1
Barrett, J. O. W. 1920 *Biochem. J.* **14**, 189
Bergell, P. & Feigl, J. 1908 *Hoppe-Seyl. Z.* **55**, 173.

Bergmann, M. 1941 *J. Biol Chem.* **139**, 481
Bergmann, M. & Niemann, C. 1938 *J. Biol. Chem.* **122**, 577.
Bernheim, F, Neurath, H. & Erickson, J O 1942 *J. Biol. Chem.* **144**, 259.
Block, R. J. 1934 *J Biol Chem* **104**, 339
Block, R. J. 1940 *J Biol Chem* **133**, 67.
Burk, N F & Greenberg, D M. 1930 *J. Biol Chem* **87**, 230.
Buswell, A. M, Krebs, K F. & Rodebush, W H. 1940 *J. Phys. Chem* **44**, 1126.
Carothers, W. H 1940 *High Polymers*, **1**, 1st ed, p 187 New York Interscience Publishers.
Chibnall, A C. 1942 *Proc Roy Soc* B, **131**, 136
Clibbens, D. A & Geake, A. 1928 *J Text Inst*, *Manchr*, **19**, T77
Clibbens, D A. & Little, A H 1936 *J. Text. Inst*, *Manchr*, **27**, T285.
Cohn, E. J. 1938 *Cold Spr Harb Symp. Quant. Biol.* **6**, 17
Cohn, E. J., Edsall, J T & Blanchard, M H. 1934 *J. Biol. Chem.* **105**, 31.
Colombo, G 1932 *Boll Staz sper. Seta, Milano*, **2**, 82.
Crammer, J L. & Neuberger, A. 1943 *Biochem J* **37**, 302
Damodaran, M. 1932 *Biochem J.* **26**, 235.
Damodaran, M., Jaaback, G. & Chibnall, A. C 1932 *Biochem. J.* **26**, 1704.
Dunn, M S, Camien, M. N, Shankman, S. & Rockland, L B. 1945 *J. Biol Chem.* **159**, 653.
Fischer, E. 1901 *Ber dtsch. chem Ges* **34**, 433.
Foà, C 1912 *Z. chem Industr. Kolloide*, **10**, 7.
Folin, O. & Malmros, H 1929 *J Biol. Chem.* **83**, 115.
Garrott, C H & Howitt, F O 1941 *J. Text. Inst*, *Manchr*, **32**, T1.
Gilson, G 1904 *Cellule*, **10**, 39.
Gordon, A H., Martin, A J P. & Synge, R L M 1941 *Biochem J* **35**, 1360
Gordon, A. H., Martin, A. J. P. & Synge, R. L M 1943 *Biochem. J* **37**, 541.
Guest, G H 1939 *Canad. J Res* **17**, *B*, 143
Harris, M. 1941 *Text. Res* **11**, 117.
Herzog, R. O 1928 *Helv. chim Acta*, **11**, 529.
Jesserer, H & Lieben, F 1937 *Biochem. Z.* **292**, 403.
Kratky, O & Kuriyama, S 1931 *Hoppe-Seyl Z* B, **11**, 363.
Kuntzel, A. & Dröscher, T 1940 *Biochem Z.* **306**, 177.
Lieser, T 1940 *Cellulose-Chem.* **18**, 121.
Lin, K. H, Wu, H. & Chen, T T. 1928 *Chin J Physiol.* **2**, 107.
Linderstrøm-Lang, K. 1939 *Proc. Roy. Soc.* B, **127**, 17.
Lloyd, D. J. & Biddor, P. B. 1935 *Trans. Faraday Soc.* **31**, 864.
Lyons, R. N. 1945 *Nature*, **155**, 633.
McMahan, J. R & Snell, E E 1944 *J. Biol Chem.* **152**, 83.
Martin, A. J. P & Synge, R L M. 1941 *Biochem J* **35**, 91.
Meyer, K H, Fuld, M. & Klemm, O 1940 *Helv. chim. Acta*, **23**, 1441.
Meyer, K H & Jeannerat, J. 1939 *Helv chim. Acta*, **22**, 22.
Mirsky, A. E. & Anson, M L 1936 *J. Gen. Physiol.* **19**, 439, 451.
Moore, E. & Stein, W. H 1943 *J. Biol Chem.* **150**, 113.
Neurath, H, Cooper, G R & Erickson, J O. 1942 *J. Biol. Chem.* **142**, 249
Neurath, H., Greenstein, J. P, Putnam, F. W. & Erickson, J. O. 1944 *Chem. Rev.* **34**, 157.
Nicolet, B. H. & Saidel, L J. 1941 *J Biol Chem.* **139**, 477.
Nicolet, B. H. & Shinn, L A 1941 *J. Biol. Chem.* **139**, 687.
Ongaro, D. 1933 *G. chim industr.* **15**, 506.
Osborne, T. B, Leavenworth, C S & Brautlecht, C. A. 1908 *Amer J Physiol.* **23**, 180.
Roche, J, Roche, A., Adair, G. S. & Adair, M E 1932 *Biochem J* **26**, 1811.
Spiegel-Adolf, M. & Henny, G C 1941 *J. Phys. Chem.* **44**, 1126.
Stokes, J. L. Gunness, M., Dwyer, I. M. & Caswell, M. C. 1945 *J. Biol. Chem.* **160**, 35.
Synge, R. L. M. 1943 *Chem. Rev.* **32**, 135.
Takamatsu, Y. 1933 *J Soc. Chem Ind. Japan*, **36**, 566.
Vickery, H B. & Block, R. J. 1931 *J. Biol. Chem.* **93**, 105.
Vickery, H. B. & Winternitz, J. K. 1944 *J. Biol Chem.* **156**, 211.
Young, E. G. 1922 *Proc. Roy Soc.* B, **93**, 15.
Zuverkalov, D. 1927 *Hoppe-Seyl. Z.* **163**, 185.

The scattering of 4·2 MeV protons by protons

By A. N. May and C F. Powell

(*Communicated by A. M. Tyndall, F.R.S —Received 25 May 1946*)

The methods described in a previous paper have been applied to an examination of the scattering of 4 2 MeV protons by protons The angular distribution of the scattered protons is found to be isotropic, in the centre-of-mass co-ordinate system of the colliding particles, from 50 to 140° in the angle of scattering. A method of determining the absolute cross-section for proton-proton scattering at different energies, depending on the simultaneous recording of the protons scattered from hydrogen and from heavy nuclei is described, and the conditions necessary for its application at higher energies are discussed

1. Introduction

In a previous paper (Chadwick, May, Pickavance & Powell 1944) which will be referred to as (I), an experimental method was described for investigating the scattering of high energy protons and deuterons generated by the cyclotron. A critical examination of the apparatus showed it to be a powerful and reliable instrument In this, and in the papers which follow, experiments are described in which the method is applied to an investigation of the scattering of 4 MeV protons and 7 MeV deuterons by ten different elements. The results were obtained in the years 1940 to 1942 their preparation for publication having been delayed by the war.

It was anticipated that if proper use were to be made of the scattering method for testing any formalism describing the interaction of nuclear particles, it would be necessary to make investigations with different elements in a large range of values of the energy of the particles composing the primary beam. These anticipations have been confirmed by the experiments to be described. Especially in the work with deuterons, a number of features of great interest and complexity in connexion with (d, p) reactions have been discovered, the elucidation of which would appear to be of importance in achieving an understanding of nuclear processes in general. It will be necessary, however, to study these phenomena with beams of particles of different energies, and, apart from certain experiments which are of immediate interest, such as the scattering of protons and deuterons by hydrogen, deuterium and helium, and the inelastic scattering of protons by different elements, the present investigations constitute only a preliminary survey of the field From this survey we are able to indicate those collision processes which, for technical or other reasons, it seems most desirable to investigate in greater detail, and the methods to be employed. This paper deals with the results of the experiments with the 4 2 MeV proton beam delivered by the Liverpool cyclotron, the measurement on the plates being made in the H H. Wills Physical Laboratory, University of Bristol. Table 1 gives details of the exposures given for the production of the plates on which the present measurements were made.

The exposures were obtained with an earlier form of the apparatus than that described in (I), the divergence of the beam being approximately 2° instead of 1°. During the exposures the mean current in the beam was of the order of 3 μA as measured at the collecting electrode, see (I), figure 2 The current in the defined beam was of the order of 1 % of that in the main beam. No attempt was made to maintain the beam current precisely constant and the exposure times and gas pressures therefore give only a rough indication of the conditions necessary for the production of satisfactory plates with the apparatus employed but, as mentioned previously, the exposure times are not critical. 'Half-tone' plates with a thickness of 70 or 100 μ were employed throughout the whole of the present series of exposures, the methods of development and fixing being similar to those previously described (Powell 1943).

TABLE 1

plate no.	scattering gas	gas pressure cm. of Hg	exposure min.
P.10	hydrogen	76	15
P.11	acetylene	15	15
P.17	methyl iodide	3·0	1
P 19	methyl iodide	1 0	0 5

In analyzing the results of the measurements we shall have to make frequent reference to the ratio of the actual scattered intensity of a group of particles, per unit solid angle, to that calculated from the Rutherford formula on the assumption that the scattering is governed only by the Coulomb forces between the interacting particles We denote this ratio, for an angle of scattering θ in the laboratory system of co-ordinates, by an element X, by the symbol $R_X^E(\theta)$, where E is the energy of the incident particles in MeV. The corresponding quantity for an angle of scattering, ϕ, in the co-ordinate system moving with the centre of mass of the colliding particles, is $R_X^E(\phi)$

For proton-proton scattering the Rutherford formula is replaced by that due to Mott, and we use the analogous symbols $M^E(\theta)$ and $M^E(\phi)$. θ and ϕ refer throughout to the laboratory and centre-of-mass co-ordinate systems respectively and $I(\theta)$, $I(\phi)$ denote the scattered intensities per unit solid angle referred to them.

2. ANGULAR DISTRIBUTION OF THE ELASTICALLY SCATTERED PROTONS

Some features of the plate taken with the camera filled with hydrogen, P 10, have already been discussed in (I). It was shown that the mean range of the protons scattered at any angle is in satisfactory agreement with theoretical expectations and that tracks due to protons scattered from impurities in the gas, which can be distinguished from those scattered from hydrogen as a consequence of their greater range, are present in appreciable numbers only at small values of θ

Intensity measurements have been made on this plate over a range of angles from 12 5 to 50°, and over a wide range of values of r (see (I), p. 16). At the larger angles, if short tracks only are counted, the small number of recoils from impurities, which

have energies only a little less than that of the incident protons, are not included. At small angles, however, this distinction is no longer possible since the range distributions of the protons scattered from hydrogen and from heavier impurities overlap.

To correct for the tracks at small angles due to the presence of the impurities the number of long tracks at 45° was counted From these observations the number to be expected at lower angles can then be calculated on the assumption that, in the range below 45°, the variation of the scattered intensity with angle bears a constant ratio to that given by the Rutherford relationship The impurities most likely to be present are oxygen and nitrogen and the experiments on the scattering by these elements, described in the following paper, show that this assumption is very nearly true in the range from 20 to 45° A more precise correction cannot be made since the nature of the impurities is unknown The correction becomes appreciable at about 30° and is of the order of 20 % at 20°

In the plates taken with the early form of the apparatus the distributions in range of the particles scattered through small angles show, in addition to the main group of particles, a 'tail', with a continuous distribution in range, due to slit scattering and other causes We have not considered it worth while to attempt to measure the intensity of the elastically scattered protons in this region since we may expect anomalies in the scattering to be relatively small at small angles.

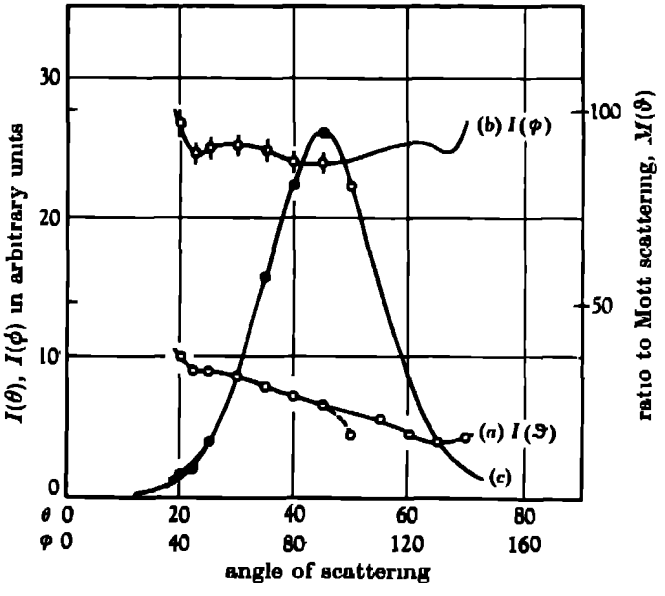


FIGURE 1

The results of the measurement are shown in figure 1, where curve (*a*) shows the intensity of the scattered protons per unit solid angle in the laboratory system of co-ordinates. The measurements of intensity become unreliable at angles of scat-

tering greater than 40° because the scattered protons lose an appreciable fraction of their energy in passing from the collision space to the photographic plate through the gas filling the camera. Their range in the emulsion then becomes so short that many of them contain only two or three grains and they are not distinguished by the observer and recorded. The scattering beyond 45° is, however, necessarily connected with that at smaller angles since the number of protons scattered between θ and $\theta+\delta\theta$ must be equal to that between $(\frac{1}{2}\pi-\theta)$ and $(\frac{1}{2}\pi-\theta)-\delta\theta$. This enables values of $I(\theta)$ at 50°, 55°, etc., to be deduced from those observed at 40°, 35°, etc. These values are shown in figure 1, curve (*a*). It will be seen that they lie on a smooth curve with those at small angles—and that the point deduced in this way for 50° is considerably higher than the observed value, because of the 'loss' of tracks in the actual measurements at this angle

The angle of scattering, ϕ, in the centre-of-mass co-ordinate system of the colliding particles is related to the corresponding angle in the laboratory system, θ, by the simple relation $\phi = 2\theta$. Similarly, if $I(\theta)$ is the scattered intensity per unit solid angle at θ and $I(\phi)$ the corresponding quantity in the centre-of-mass system at ϕ,

$$I(\phi) = \tfrac{1}{4}I(\theta)\sec\theta.$$

$I(\phi)$ can therefore be determined from the observed values of $I(\theta)$ and the results are shown in figure 1, curve (*b*) The vertical lines give the r m.s. error deduced in the usual way from the number of tracks counted. It will be seen that the scattered intensity $I(\phi)$, in the interval from 50 to 90°, i.e. in the region where the Coulomb scattering is very small compared with the anomalous scattering, is nearly constant. The departures of the observations from a line $I(\phi)$ = constant, corresponding to isotropic scattering, are of the order of only 2 % of the mean intensity, and although this is greater than the standard deviations associated with the number of tracks counted, the departure from isotropy cannot be regarded as significant. We therefore conclude that, at these angles, the observations correspond to pure '*s*' scattering within the limits set by experimental errors Creutz & Wilson (1942) have reported that proton-proton scattering is isotropic at 8 MeV, apart from the Coulomb scattering at small angles, and it is clear that more energetic proton beams will be needed to give any easily detectable effects from '*p*' and other waves of higher order.

3 Absolute cross-section for proton-proton scattering

The determination of the absolute cross-section for this process, in addition to the variation of the intensity of the scattered protons with angle, is of great importance for the determination of the character and magnitude of the forces between the two particles. For such a determination we require to know the total number of primary protons passing down the scattering tube during the exposure. In our first experiments there was no provision for determining the current in the defined beam during the exposure and we therefore adopted an alternative method.

For this purpose the intensities of the protons recoiling from hydrogen and from some heavier type of nucleus were measured at a particular angle of scattering. The ratio of the numbers of atoms of the two types in the scattering gas can be defined by employing a suitable gaseous compound or a mixture of gases in known proportions At angles of scattering in the neighbourhood of 45° the difference in energy of the two groups of scattered protons is sufficiently great to ensure that the range distribution shows them clearly resolved and the ratio of their intensities can thus be determined. If, in the conditions of the experiment, the intensity of the scattering from the heavy nuclei is given by the Rutherford relationship, then the total flux of primary protons during the exposure can be calculated and hence the cross-section for the proton-proton scattering process at the same angle

As a preliminary test of the method we examined the proton recoils at 45° obtained with methyl iodide and acetylene as the scattering gases.

4. Experiments with methyl iodide

A typical example of the distribution in energy of the particles scattered at 45° in these exposures is shown in figure 2, where the two groups are seen to be clearly resolved It will be seen that the 'energy-width' at half maximum of the protons scattered by the iodine nuclei is considerably less than that for the proton-proton group The increased width of this latter group is easily explained by the known divergence of the primary beam, $\pm 2°$. Thus the energy, E, of the scattered protons is given by the relation $E = E_0 \cos^2\theta$ Allowing $\pm 0{\cdot}5°$ for the small angle deviations of the particles in traversing the window and the gas, before reaching the plate, the particles which enter the emulsion at points corresponding to 45° will have been scattered from the primary beam through angles varying from 42·5° to 47·5°. Putting $E_0 = 4{\cdot}2$ MeV, E ranges from 2·38 MeV at 42·5° to 1·92 MeV at 47 5°; a difference of 0 46 MeV, in good agreement with the observations. Thus even if the primary beam is perfectly homogeneous in energy the energy-width of the proton-proton peak at 45° will be 0 5 MeV in the geometrical conditions provided by the early form of the apparatus The energy of the protons scattered by the iodine varies very little, however, with the angle of scattering and the observed width agrees with that obtained with homogeneous groups of protons of a similar energy in previous experiments (Powell 1943). We may thus conclude that any variation in energy of the protons in the primary beam is small compared with the energy resolution given by the method. The results emphasize the advantages of a more closely defined beam, as produced in the later form of apparatus, where the divergence is reduced to $\pm 1°$

The number of tracks in the two groups of particles in a given area of the plate gives, within the limits of statistical error, the ratio of the scattering from $3H$ and from I at the chosen angle of scattering, 45°. A small correction must be applied for the contribution to the long tracks of the protons scattered by the carbon nuclei which are not resolved from those from the iodine. This correction can be deduced

from the work on the acetylene plate which gives the ratio of the scattering from carbon and from hydrogen at the same angle.

A second correction must be applied to the observed number of particles in the shorter-range group since a certain proportion of them must be attributed to the 'tail' of the longer group. The main group in the range distribution at a given angle of scattering, produced by a homogeneous group of particles, is accompanied by a low intensity tail, uniformly distributed in range, due to the penetration of the edges

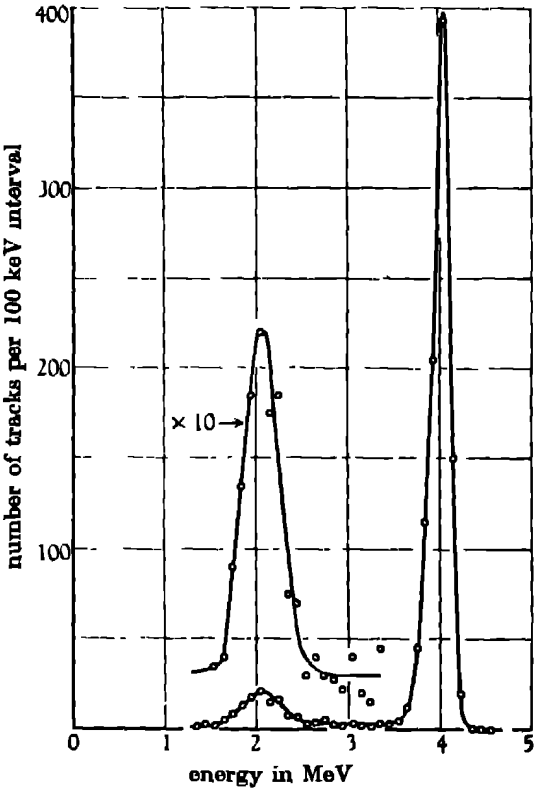


FIGURE 2

of the defining slits by a small fraction of the scattered particles and a corresponding reduction in their energy, see (I), p 23. Figure 3 shows a typical range distribution, from measurements on the methyl iodide plates, in which the continuous distribution of tracks between the two peaks will be seen. In this distribution the number of tracks in the range from 6 to 10 divisions is 53 and in the short range group, from 1·2 to 5·2 divisions, 238. We therefore assume that, in this case, 185 tracks are to be attributed to proton-proton scattering

Altogether about 4800 tracks were measured on the methyl iodide plates of which 4204 were long and 397 were from proton-proton scattering The carbon is estimated to contribute 93 long tracks to the group of longer range so that we deduce for

the ratio of the intensity of the scattering from hydrogen and from iodine at 45° the value:

$$\frac{3(H)}{I} = \frac{397}{4111} = (9{\cdot}65 \pm 0{\cdot}8) \times 10^{-2}.$$

Now according to the formulae of Mott and of Rutherford this quantity should have a value at 45° of $1{\cdot}03 \times 10^{-2}$. Assuming that the scattering from iodine is accurately given by the Rutherford relationship we therefore conclude that the quantity, $M^{4\cdot2}(45°)$, equals 94 ± 6.

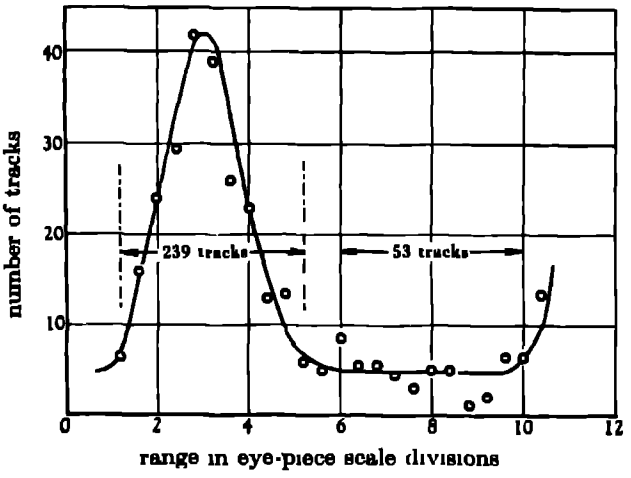


FIGURE 3

5. EXPERIMENTS WITH ACETYLENE

Similar observations have been made using the plates obtained with acetylene as the scattering gas. An investigation was made of the intensity of the long and short tracks at a number of angles in the interval from 25 to 40° in the angle of scattering. The lengths of 2600 tracks were measured and examples of the observed distribution in range of the tracks at four different angles are shown in figure 4. Figure 5 shows an example of the results transformed to an energy distribution, and it will be seen that, in this case also, the width of the peak due to proton-proton scattering is greater than that due to scattering from the heavier nuclei.

The variation with angle of the intensity of the long tracks, due to the elastic scattering of the primary protons by carbon nuclei, is shown in figure 6. In this diagram, curve (a) shows the variation with angle of the scattered intensity and curve (b) the same results expressed in the form of the ratio, $R(\theta)$. It will be seen that $R(\theta)$ is nearly constant in the range of angles from 20 to 45° to within the accuracy of the experiment. If we assume that the absolute intensity is also given by the Rutherford equation in this range of angles, i.e. that $R_C^{4\cdot2}(\theta) = 1$, then a comparison of the number of long and short tracks at 45° enables us to obtain a second estimate for the quantity M (45°). The value so obtained is $M^{4\cdot2}(45°) = 72 \pm 7$.

It is clear, from a comparison of this result with that obtained with methyl iodide, that the cross-section for the scattering of 4·2 MeV protons by carbon, at these angles, is different from the value given by the Rutherford formula. It is shown in the following paper that there are good reasons for anticipating such a result, even at

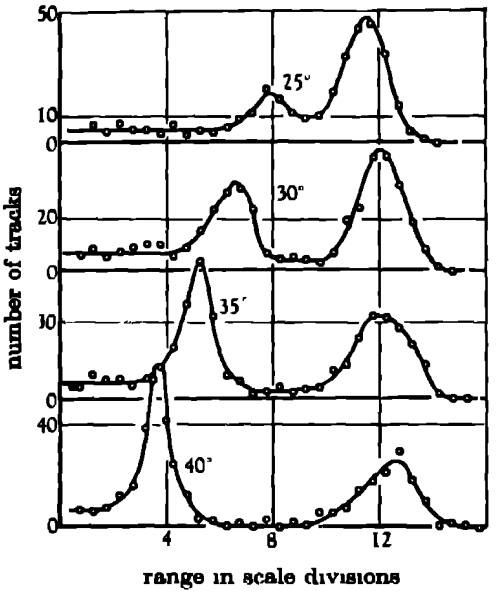


FIGURE 4

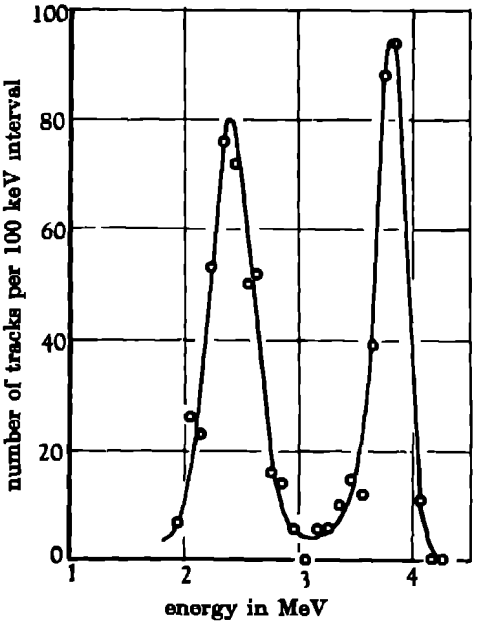


FIGURE 5

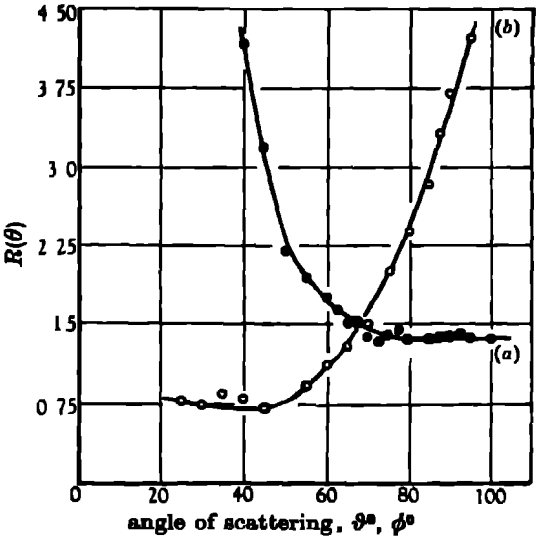


FIGURE 6, Elastic scattering of 4·2 MeV protons by carbon
(*a*) I (ϑ) arbitrary units, (*b*) $R(\theta)$.

angles of scattering in the interval from 20 to 45° The value of R (45°) is easily found by a comparison of the results of the experiments on methyl iodide and acetylene at 45° and we obtain the result

$$R_C^{4\cdot2}(45°) = \tfrac{7}{64} = 0\cdot76 \pm 0\cdot08.$$

The scale of the results of curve (b) in figure 6 has been chosen to correspond with this value of $R(\theta)$ at 45°. We have already pointed out that there was no stringent control of the purity of the target gases. The presence of a considerable pressure of air would not, however, seriously affect the result The most likely impurity, in the experiments with methyl iodide, is water vapour from the emulsion which, by contributing hydrogen to the scattering gas, would tend to make the observed value of M too high The values obtained from the two plates P.17 and P 19 agree, however, within the limits of statistical error and this suggests that the disturbing effects due to water vapour were in fact small The methyl iodide pressure was varied by a factor of three in the two exposures and the water vapour pressure would be unlikely to change by the same ratio in the two cases.

6. Discussion

(a) *Absolute cross-section*

We have obtained the value 94 ± 6 for the quantity $M^{4\cdot2}(45°)$. Breit, Thaxton & Eisenbud (1939) have calculated this ratio for pure 's' scattering, for protons of different energies, using various forms of 'potential well' to represent the forces at close distances of approach of the two protons The values they give for the ratio M (45°) for 4 MeV protons, range up to 95 for the narrowest and deepest wells and our result corresponds to a value for the phase change K in the wave of zero angular momentum of $54 \pm 2\cdot5°$. Ragane, Kanne & Taschek (1941) have shown that the results obtained with slow protons in the energy range from 200 to 300 eKv are in agreement with a 'well' of radius e^2/mc^2 and a depth of 10 52 MeV. This well gives $K = 52°$ at 4 MeV and a value of M (45°) of 85 Our results are therefore in accord, to within the errors of the measurements, with the values predicted by this model In a later publication, Landau & Smorodinsky (1944) give a formula for the phase shift, K_0, in terms of a constant, α_{pp}, which represents the logarithmic derivative of the wave function at the origin. They give an empirical formula for α_{pp} which fits the results of previous investigators extending up to 2 4 MeV. From this formula we calculated the phase shift at 4 MeV and obtained the value $K = 53°$. We therefore conclude that our observations show no significant departures from the values calculated from theories based on previous experiments at lower energies. The importance of work with beams of protons of higher energy is thus emphasized.

Using the value $K = 54°$ deduced from the observed value of M at 45°, we have calculated this ratio for other angles. The results are shown in the smooth curve of figure 1, curve (c). It will be seen that there is good agreement with the experimental results except in the region of small angles where the effects due to slit scattering begin to become appreciable.

(*b*) *Appraisal of the method*

From the experience gained in the course of the present experiments we can estimate the usefulness of the method for the examination of the proton-proton collision and the steps which must be taken to improve it. We have already pointed out the importance of extending the experiments to higher energies and our experience with fast protons, which extends up to 13 MeV, indicates that the method could be successfully applied with primary particles of energy up to 20 to 30 MeV Such beams would produce 10 to 15 MeV protons at 45°, and these can be recorded by using plates with an emulsion thickness of 100μ and an angle of approach of the scattered protons to the plate of 5°

The method here employed for the determination of the values of the absolute cross-section would seem to give a useful check on the usual method based on the direct determination of the total flux of charge in the primary beam during the exposure. The use of chemical compounds in the form of vapours or gases ensures that the relative numbers of scattering nuclei of different types are known with great precision, provided that adequate control of purity is established. In future work it would appear desirable to employ gaseous mixtures of known composition in such proportions that the numbers of protons in the groups scattered from the two types of nuclei are approximately equal.

Using the ordinary half-tone emulsions the counting and measuring of the tracks is sufficiently fast under good conditions to make it possible to reduce the errors in the determination of the scattered intensity at any angle, due to statistical fluctuations, to the order of 2 % in the course of 4 days work by a single observer. Further, the improved registration which can now be obtained with concentrated half-tone emulsions leads to greatly improved methods of counting It thus becomes possible to contemplate experiments in which much greater numbers of particles are counted and measured so that substantial reductions in the standard deviations, and corresponding improvements in the accuracy of the results, are obtained.

A second improvement in the method, suggested by the present results, is to introduce adequate control of the purity of the gas in the collision space by eliminating the presence of water vapour from the gelatine This can be easily achieved by making the collision space gas tight, a modification which is also an essential preliminary to making experiments with rare gaseous isotopes, available only in small quantities, when it is important to recover the gas after the exposure Although the windows in the present form of the apparatus are of relatively large area, we have found it possible to cover them with foils, with a stopping power of only a few mm of air, which are able to withstand pressure differences of 10 cm of mercury.

The experimental work described in this and the two following papers was undertaken, as part of an extended programme, in association with Sir James Chadwick and his collaborators in Liverpool. Because of the war we alone were able to make measurements on the plates and we are indebted to them for their agreement to our publication of the results.

We are greatly indebted to Professor A M. Tyndall for the facilities which made it possible for us to analyze the plates in the H. H. Wills Physical Laboratory of the University of Bristol.

References

Breit, Thaxton & Eisenbud 1939 *Phys. Rev.* **55**, 1018.
Chadwick, May, Pickavance & Powell 1944 *Proc Roy. Soc.* A, **183**, 1.
Creutz & Wilson 1942 *Phys Rev* **61**, 388
Landau & Smorodinsky 1944 *J. Phys. U.S S R.* **8**, 154.
Powell 1943 *Proc. Roy Soc.* A, **181**, 344
Ragane, Kanne & Taschek 1941 *Phys. Rev.* **61**, 628.
Wilson & Creutz 1941 *Phys. Rev.* **59**, 806.

The scattering of 4·2 MeV protons by deuterium, helium and other light elements

By H. Heitler, A. N. May and C. F. Powell

(*Communicated by A. M. Tyndall, F.R S —Received* 4 *June* 1946)

The methods previously described have been applied to a study of the elastic scattering of 4 2 MeV protons by deuterium, helium, nitrogen, oxygen, neon, argon, chlorine and bromine. The results obtained with deuterium and helium as the scattering gases are compared with those of other experimenters working with protons of different energies The angular distributions of the protons elastically scattered by nitrogen and oxygen, together with the results for carbon obtained in previous measurements, can be simply described in terms of an S-component interacting with the Coulomb term in the formula for the scattered wave Values are obtained in each case for the magnitude of the parameter which defines the amplitude and phase of the S-wave The angular distribution of the protons inelastically scattered by neon has been investigated and evidence obtained for similar collision processes in experiments with argon and chlorine.

1. Introduction

In continuation of previous experiments involving the application of the photographic method to problems in nuclear physics, we have made observations on the scattering of 4·2 MeV protons by gas targets of deuterium, helium and six other light elements. The essential features of the method have been described in two previous papers (Chadwick, May, Pickavance & Powell 1944; May & Powell 1947 (see p. 170)) which will be referred to as I and II respectively. The proton beam was obtained from the Liverpool cyclotron, the examination of the plates being carried out in the H. H Wills Physical Laboratory, University of Bristol, in the years 1940 and 1941.

Table 1 gives details of the exposures given to the plates on which the present measurements were made.

TABLE 1

plate number	gas	pressure (cm. of Hg)	exposure (min.)
	exposures with 4·2 MeV protons		
P 20	deuterium	13·5	10
P 18	helium	10·0	4
P 21	oxygen	10·0	3
P 7	nitrogen	10·0	10
P 9	neon	5·5	15
P 22	argon	5·0	5
P 16	carbon tetrachloride	2·5	3
P 15	bromoform	0·5	2
	exposures with 6·7 MeV deuterons		
D 11	hydrogen	76·0	3

It will be seen from the description of the experimental method given in I that the particles entering a given small area of the plate have all been scattered from the primary beam through angles differing by less than 2° from a certain mean value, the angle of scattering, θ. This angle is determined by the position of the area under consideration relative to the scattering gap, and with the later forms of the apparatus, such as is described in I, observations can be made at points in the plate corresponding to any value of θ from 10 to 170°. The plates in series P were obtained, however, with an earlier form of the apparatus in which the particles scattered through angles greater than 100° were not recorded and the present observations are therefore confined to the restricted interval in the value of θ from 15 to 100°. The distribution in range of the particles at any angle of scattering is determined by measuring the length of the tracks which they produce in the emulsion and the intensity of the scattered beam is given by counts of the number of tracks per unit area, both observations being carried out at the appropriate region in the plate.

The particles entering the plate in any given element of its surface are all produced by collisions with the nuclei of gas atoms in a small volume traversed by the primary beam. We refer to the 'centre of mass' of the points of origin of these scattered particles as the 'mean point of scattering' for the particular value of θ which corresponds to the position of the element of area under consideration. It is a feature of the method that the mean energy of the primary protons at the mean point of scattering, E_0^θ, varies with θ when using gaseous 'targets'. This is due to the retardation of the primary particles by the gas filling the camera and to the change with θ of the mean point of scattering; see I, p. 23. In the present experiments, however, this energy change is small, in the range of angles at which measurements can be made, because of the low gas pressures employed in the camera. Thus in the most unfavourable case, that of oxygen, plate P 21, E_0^θ changes by only 20 keV, as θ increases from 20 to 90°, a value much less than the spread in energy of the primary particles delivered by the cyclotron.

2 Notation

It is convenient to summarize the notation employed in this and the following paper in order to avoid unnecessary repetition

θ, ϕ	angle of scattering of a primary particle in the laboratory and centre-of-mass systems of co-ordinates respectively.
η	angle of projection of a target nucleus in the laboratory system.
$I(\theta)$, $I(\eta)$, $I(\phi)$	scattered or projected intensity per unit solid angle in the corresponding co-ordinate systems.
E_0^m, ρ_0^m	the mean energy and mean range in standard air, respectively, of the primary particles emerging through the mica window from the cyclotron tank.
E_0^θ, ρ_0^θ	the mean energy and range, respectively, of the primary particles at the mean point of scattering of the particles observed at θ
E_s^θ, ρ_s^θ	the mean energy and range, respectively, of the particles scattered through the angle θ.
$R(\theta)$, $R(\phi)$	the ratio of the observed scattered intensity at θ or ϕ to the value calculated from the Rutherford formula.
$M(\theta)$, $M(\phi)$	a similar ratio to the above for the particular case of collision between indistinguishable particles—the ratio of the observed intensity to that given by the Mott formula.

In discussing the collision processes frequent reference is made to a system of co-ordinates moving with the centre of mass of the interacting particles or to a system at rest relative to the laboratory.

3 The proton-deuteron collision

The results obtained in experiments on the scattering of 4·2 MeV protons by deuterium and of 0·7 MeV deuterons by hydrogen may be considered together since the two types of collision are similar in the centre-of-mass system of co-ordinates. In this system the proton and deuteron approach one another with equal and opposite momentum and each is scattered through the same angle, ϕ. $I(\phi)$, the scattered intensity in the centre-of-mass system can thus be determined either by observations on the intensity of the scattered primary particles, $I(\theta)$, or of the projected nuclei, $I(\eta)$

(*a*) *Scattering of protons by deuterium, plate P* 20

The results obtained in measurements on this plate were discussed in I where it was shown that, owing to the presence of impurities in the deuterium gas employed, the scattered beam contained protons elastically scattered from ordinary hydrogen and heavier nuclei in addition to those from deuterium. This is illustrated in figure 1

which shows the results of measurements of the distribution in range of the scattered particles at $\theta = 45°$, the arrows indicating the scattering nuclei responsible for the various peaks.

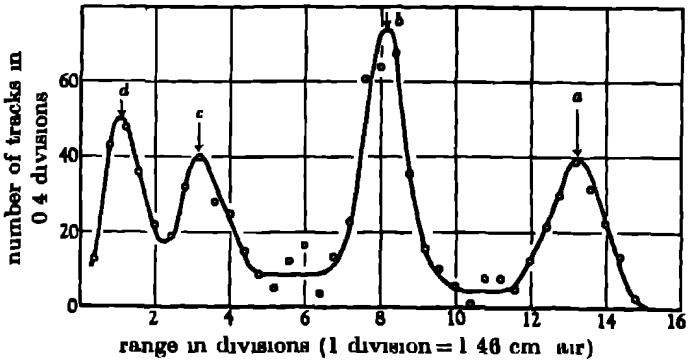


FIGURE 1 Distribution in range of the particles scattered at 45° from impure deuterium, protons elastically scattered from oxygen and nitrogen, peak (*a*); from deuterium, peak (*b*), from hydrogen, peak (*c*). Peak (*d*) is due to deuterons projected by the primary protons. The observed range is not corrected for the stopping power of the gas and foil

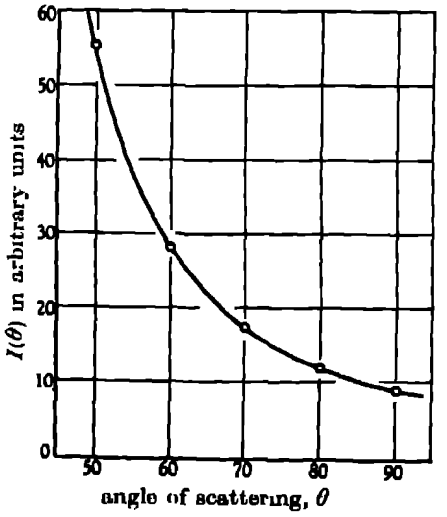


FIGURE 2. Angular distribution of 4 2 protons scattered from deuterium

At angles of scattering between 40 and 90° the protons scattered by deuterium can be clearly distinguished from other groups and their intensity determined Beyond 90° the energy of these particles is so low that the tracks which they produce on entering the emulsion contain only a few grains and the counts on which the intensity measurements are based become unreliable, whilst at angles less than 40° they cannot be distinguished from the protons scattered by nitrogen and oxygen. The intensity measurements are therefore confined to a restricted range in the value of θ and the results are shown in figure 2.

The deuterons projected by the impact of the primary protons can be detected at angles between $\eta = 30°$ and $\eta = 45°$ but their range is too low and too near that of the group arising from p-p scattering for reliable estimates of their intensity to be made. In more favourable conditions, using pure deuterium and primary particles of higher energy, such observations would be useful in supplementing the measurements on the scattered protons. The projection of a deuteron corresponds, for small values of η, to the backward scattering of a primary proton and measurements of the intensity of such projected particles lead to determinations of $I(\phi)$ at large values of ϕ. It is difficult to obtain such information from observations on the scattered protons because of their very low energy at large angles of scattering

(*b*) *Scattering of deuterons by hydrogen, plate D* 11

The measurements on this plate have also been referred to in I, p. 19 In elastic collisions with hydrogen the deuterons cannot be scattered through angles greater than 30° and the form of the curve representing the angular variation of the scattered intensity shows a sharp edge at this angle which was employed to demonstrate the angular resolving power of the apparatus

Simple considerations show that at any angle of scattering less than 30° there will be two homogeneous groups of elastically scattered deuterons of different energy Thus let v be the velocity of the incident deuterons In a collision with a proton the velocity of the centre of mass is $2v/3$ and, in this co-ordinate system, the velocities of the deuteron and the proton are $v/3$ and $2v/3$ respectively. Reference to figure 3 shows that if the angle of scattering of the two particles in the centre of mass system is ϕ, this angle is related to the corresponding angle of scattering of the deuteron in the laboratory co-ordinate system, θ by the equation

$$2 \sin \theta = \sin (\phi - \theta).$$

There are two values of ϕ between 0 and π, ϕ_1 and ϕ_2, which correspond to any given value of θ less than 30°, the corresponding velocities of the scattered deuterons being represented in figure 3 by OB and OA respectively. Similarly the projection of a proton at an angle η corresponds to it being scattered in the centre of mass system, through an angle ϕ, where $\phi = \pi - 2\eta$.

In addition to the projected protons, we observe in the conditions of our exposure only the groups of deuterons of longer range, except in the region near 30° The energy of the slower particles diminishes rapidly as we proceed to smaller angles so that they are unable to traverse the foil and the gas filling the camera to reach the photographic plate We therefore measure the intensity of the projected protons and of the faster group of the elastically scattered deuterons and the results are shown in paper I, figure 12.

(*c*) *Comparison of the p-d and d-p scattering*

To compare the results of the two experiments we transform the observations to the centre-of-mass co-ordinate system in each case. As we have seen, in this system the proton and deuteron approach in opposite directions and both are scattered

through the same angle ϕ. For the p-d collision the angle of scattering of the proton in the laboratory system, θ_p, can easily be expressed in terms of ϕ Thus

$$\sin\theta_p = 2\sin(\phi - \theta_p)$$

and if $I(\phi)$, $I(\theta_p)$ are the corresponding intensities per unit solid angle in the two co-ordinate systems, we have

$$I(\phi) = I(\theta_p)\frac{\sin\theta_p}{\sin\phi}\frac{d\theta_p}{d\phi}.$$

For the d-p collision, we have seen that the angle η at which the proton is projected in the laboratory system is given by $\phi = \pi - 2\eta$ and that values of $I(\phi)$ can be determined from the observed values of $I(\eta)$. Other values of $I(\phi)$ can be deduced from the intensity, $I(\theta_d)$ of the elastically scattered deuterons The results obtained by the two methods overlap in a certain range of values of ϕ and the points representing the observations depart from a smooth curve drawn through them by amounts which correspond to the standard deviations associated with the finite number of tracks counted at any angle.

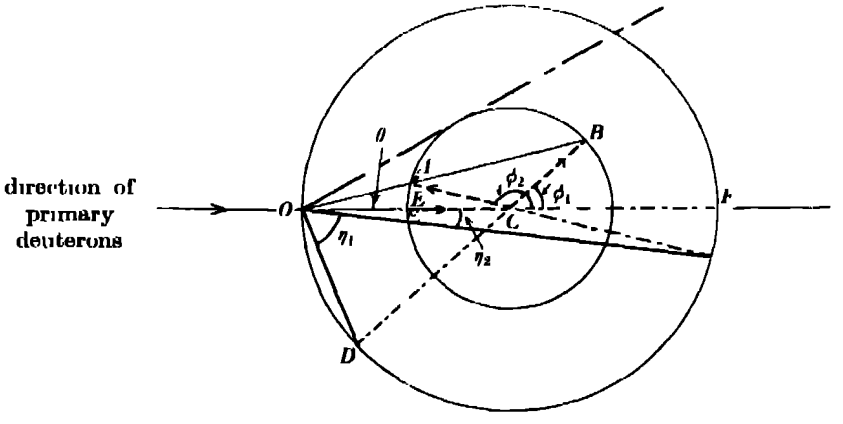


FIGURE 3. Velocity diagram for the d-p collision OC $(=2v/3)$ represents the velocity of the centre of mass of the two interacting particles relative to the laboratory In the centre-of-mass co-ordinate system the deuteron with velocity EC $(=v/3)$ approaches the proton of velocity $FC = 2v/3$ For an angle of scattering ϕ, in the centre-of-mass system the corresponding observed angle of scattering is θ, the observed velocity of the scattered deuteron being represented by the vector sum of $OC + CB = OB$ The corresponding angle of projection of the proton is η_1, its velocity being represented by OD.

The two sets of values of $I(\phi)$ obtained from the two separate experiments, are plotted against ϕ in figure 4. It is to be noticed that they are in different arbitrary units, and that the energies of the colliding particles in the centre-of-mass system are not the same Thus the collision of a 4·2 MeV proton with a deuteron is equivalent, in the centre-of-mass co-ordinate system, to the impact of a 8 4 MeV deuteron with a proton, whilst the actual energy of the deuterons employed in the present experiments is 6·7 MeV. In the case of the p-d collision, the measurements beyond $\theta = 110°$

are based on counts of protons of very short range and the fall in the curve is to be attributed to the fact that some of the particles fail to reach the plate. It may be noticed that measurements of the absolute cross-sections for the collisions could be obtained, as in the experiments on p-p scattering described in II, from exposures employing suitable gaseous compounds or mixtures

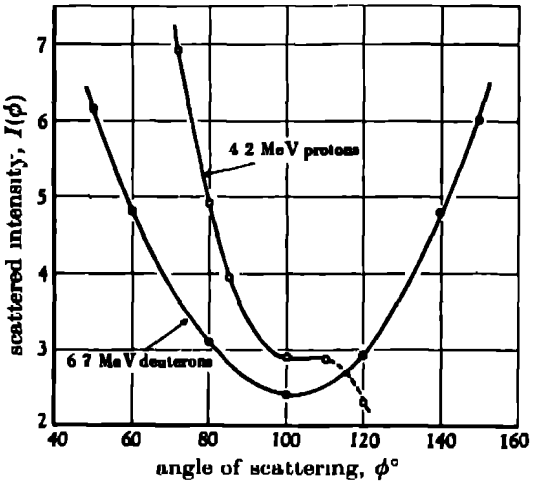


FIGURE 4. The p-d collision, variation of the scattered intensity in the centre-of-mass co-ordinate system.

In the d-p collision there are other possible processes such as that represented by the equation $H_1^1 + H_1^2 \rightarrow 2H_1^1 + n_0^1$; $Q = -2{\cdot}2$ MeV Any particles produced in this reaction will be emitted within a cone of semi-vertical angle 13·5° about the axis of the deuteron beam and, in the conditions in which our exposures were made, they would not have reached the photographic plate because of their restricted range. Similar conclusions apply to the particles from other possible reactions

Comparison with other experiments

The curves shown in figure 4 display a pronounced minimum at angles of scattering near $\phi = 100°$, and the scattered intensity rises sharply on each side of this minimum in a way which corresponds to a strong component proportional to $P_1(\cos\phi)$ in the amplitude of the scattered wave Similar conclusions were reached by other experimenters using primary protons of much lower energy Thus Taschek (1942) and Tuve, Heydenburg & Hafstad (1936), who made experiments with 200 and 800 keV protons respectively, concluded that the observed angular variation of the intensity of the scattered particles could not be accounted for in terms of an s component in the scattered wave alone, and that it was necessary to assume appreciable contributions from waves of higher angular momentum.

There appears to be no detailed theoretical treatment of the proton-deuteron collision, although Buckingham & Massey (1941) have dealt with the similar problem

of the scattering of neutrons by deuterons. To provide sufficient material for testing any formalism describing the interaction of the two particles it would be of great advantage if scattering experiments could be made at a number of values of the energy of the primary particles in a wide range.

4. The proton-helion collision

The plate obtained in the exposure with helium as the scattering gas, P 18, was very satisfactory and measurements of the range and intensity of the scattered particles were made at a large number of angles in the interval from $\theta = 15°$ to $\theta = 105°$ The energy of the scattered particles was found to vary with θ in the way to be expected if they are produced by the elastic scattering of the primary protons by helium nuclei The range of the α-particles projected by proton collision was too low for them to reach the plate in the conditions of our exposure

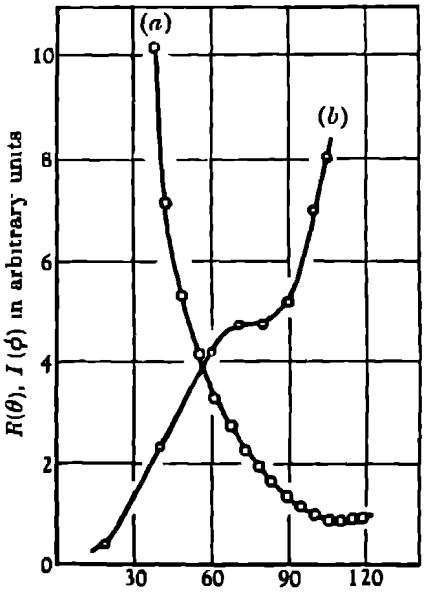


Figure 5. Scattering of 4 2 MeV protons by helium.

The results of the intensity measurements are given in figure 5 (*a*) where they are shown transformed to the centre-of-mass sytem of co-ordinates The results are also expressed in terms of the quantity $R(\theta)$, the ratio of the actual scattered intensity to that expected from the Rutherford formula. Since these results were obtained, Heydenburg & Ramsey (1941) have published an account of observations on the p-α collision, working with protons in the energy interval between 1 and 3 MeV. If we plot their values for $R(\theta)$ at a particular angle as a function of the proton energy

we can extrapolate to obtain the absolute value of R at this angle for 4·2 MeV protons. From our present results we can then determine the values of R at 4·2 MeV for all other angles. This has been done in figure 6 where Heydenburg and Ramsey's measurements are represented by open circles It will be seen that the value $R(60°) = 51$ at 4·2 MeV, lies on the curve through Heydenburg and Ramsey's results and that the values at other angles thus determined also lie smoothly on their curves

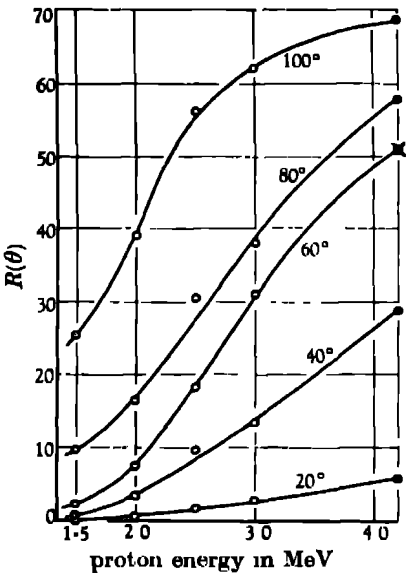


FIGURE 6. Ratio to Rutherford scattering in the p-α collision. The full circles represent the results of the present measurements adjusted to give a value for R (60°) of 51 at 4 2 MeV. The results of the measurements of Heydenburg and Ramsey are shown by the open circles

M. Tsien informs us in private discussion that he has made a theoretical investigation of the p-α collision, treating it as a 'mixed-resonance' problem His results are in satisfactory agreement with the measurements shown in figure 5 and with his own observations on the scattering of α-particles by protons. Landau & Smorodinsky (1944) point out that, in order to provide sufficient information for a rigid check of any theoretical interpretation of the collision, it is important to make a series of experiments at closely spaced intervals in the energy of the primary particles.

5. EXPERIMENTS WITH OXYGEN AND NITROGEN

For investigating the scattering of protons by the nuclei N^{14} and O^{16}, exposures were made with the camera filled in turn with commercial nitrogen and oxygen from cylinders. In both cases the range distribution of the scattered particles shows a single peak with a mean range at any angle equal to that expected for elastically

scattered protons. The results for the two gases are shown in figure 7 and 8 which give the observed values of $I(\theta)$ for the scattering of 4·2 MeV protons from nitrogen and oxygen respectively.

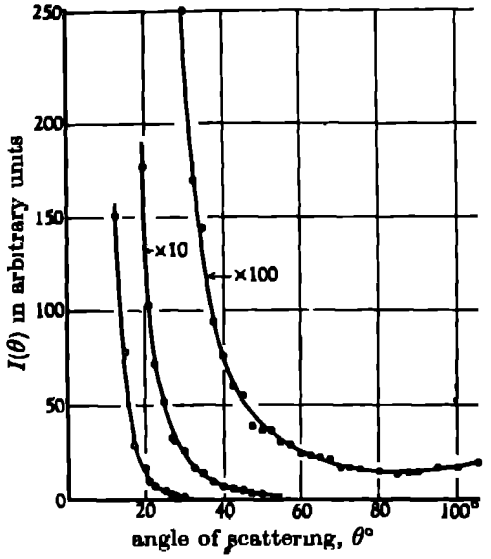


FIGURE 7. Elastic scattering of 4·2 MeV protons by nitrogen.

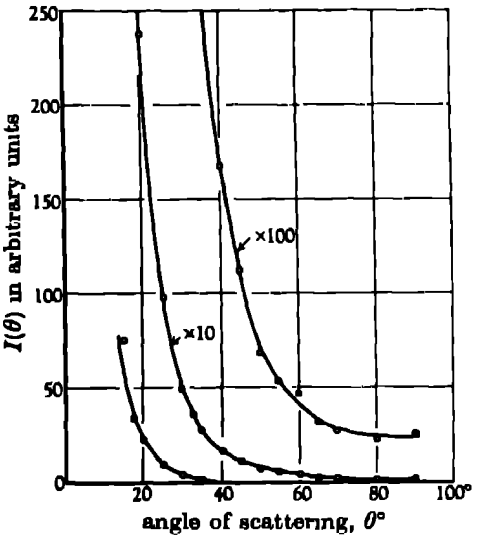


FIGURE 8. Elastic scattering of 4·2 MeV protons by oxygen.

The main features of these curves are well accounted for in terms of a formula (Mott & Massey 1933) which expresses the scattered intensity in terms of the interaction of the Coulomb term with S, P waves, etc. This formula may be expressed in the notation of the present paper in the form

$$R(\phi) = \left| 1 + \frac{i}{\alpha} \sin^2 \tfrac{1}{2}\phi e^{i\alpha \log \sin^2 \frac{1}{2}\phi} \sum_{n=0}^{\infty} (2n+1)\{e^{2i\xi_n} - 1\} P_n(\cos\phi) \right|^2,$$

where

$$\alpha = \frac{2\pi Z' Z'' e^2}{hv}$$

Z' being the atomic number of the incident particles, Z'' that of the target nuclei and v the velocity of the primary particles in the centre-of-mass system of coordinates. The quantities ξ_0, ξ_1, etc , in this formula define both the amplitude and the phase of the different harmonic components interacting with the Coulomb term to give the amplitude of the scattered wave

We are indebted to Professor Mott and Miss E. M. Littleton for a number of numerical calculations of values of $R(\phi)$ from this formula of which examples are represented graphically in figure 9 The different curves are calculated for the values of ξ_0 indicated in the figure, the coefficients ξ_1, ξ_2, etc., corresponding to waves of higher angular momentum, being put equal to zero, and α being taken as 0·63 throughout, a value appropriate to the collision of 4·2 MeV protons with O^{16} nuclei.

A comparison of figures 7, 8 and 9 shows that there is a general similarity between the form of the curves representing the angular distribution of the scattered intensity as determined by experiment and the theoretical curves in the same angular range corresponding to a suitable value of ξ_0. It is shown in the following paper that the presence of even small contributions from waves of higher angular momentum greatly modifies the form of the curves representing angular distributions We may therefore conclude from the present results that the main features of the scattering in these collision processes can be described by the interaction of the Coulomb term with an S-wave only and that effects due to P, D waves are small

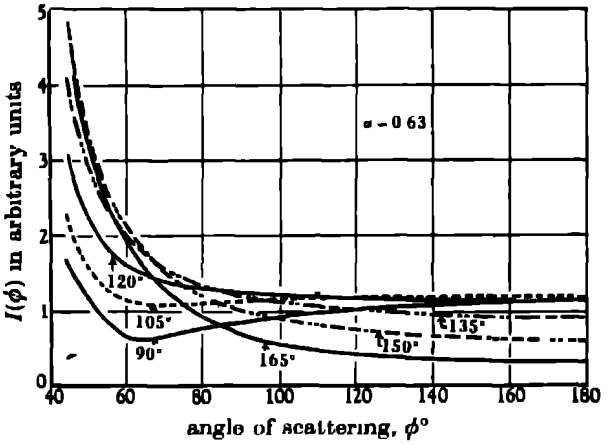


FIGURE 9. Calculated values of the variation with angle of the intensity of elastically scattered protons The value of the constant α, 0 63, corresponds to the case of the collision of 4 2 MeV protons with oxygen. The values of ξ_0, the parameter which defines the amplitude and phase of the S-wave, are indicated against the corresponding curves

To make a more precise comparison of the theoretical and experimental results it is convenient, since we cannot determine the absolute values of $R(\theta)$ from our observations, to compare the ratio of $R(\phi)$, corresponding to different values of ϕ, to the value of the same quantity at an arbitrarily chosen angle, $\phi = 60°$. Thus figure 10 shows the variation with ϕ of the quantity $R(\phi)/R\ (60°)$ as deduced from the formula of Mott and Massey for various values of ξ_0. The corresponding experimental curves for the scattering of 4·2 MeV protons by carbon, nitrogen and oxygen, are shown by the full lines in the same figure. It will be seen that in each case there is a general agreement with the theoretical curves if ξ_0 is given a suitable value. We can thus obtain approximate values for ξ_0 in the different cases which are shown in table 2.

In future experiments it will be necessary to measure the absolute values of the cross-sections for the collision processes at a number of values of the energy of the primary protons and to extend the measurements over the whole angular range. Observations at large angles of scattering are particularly important in the determination of more precise values of ξ_0 and in estimating the magnitude of contributions from higher terms.

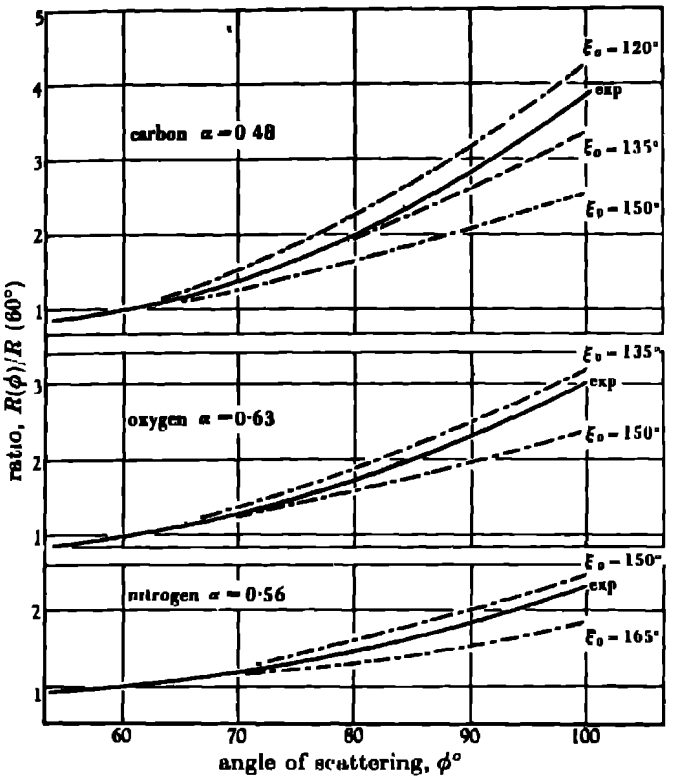


FIGURE 10 Comparison of experimental determinations of $R(\phi)/R$ (60°) with theoretical values deduced from the formula of Mott & Massey. The values of α correspond to the collision of 4 2 protons with the different nuclei.

TABLE 2. ELASTIC SCATTERING OF 4·2 MeV PROTONS

scattering nucleus	α	ξ_0
C	0·48	125°
N	0 56	155°
O	0 63	140°

6 EXPERIMENTS WITH NEON

The results of measurements on the distribution in range of the particles scattered through 90° by neon have already been published (Powell, May, Chadwick & Pickavance 1940) and show that two groups of particles are present. The more energetic of these has a mean range corresponding to that expected for protons scattered elastically We attribute the second to inelastically scattered protons from Ne^{20}, the struck nucleus being left in an excited state at 1·5 MeV after the collision. Such an excited state has already been observed in experiments on the neutrons from the reaction $F^{19}(d, n)Ne^{20}$; Bonner (1940), Powell (1940, 1943).

At smaller angles of scattering other groups of particles of low range appear, and figure 11 shows the distribution obtained by plotting together the results of all measurements made at 50, 42·5 and 27·5° The peak at 8 8 divisions corresponds to the formation of Ne^{20} in the first excited state at 1·5 MeV It seems reasonable to attribute the group at 13·0 divisions, because of its low intensity, to inelastic collisions of the primary protons with Ne^{22}, and, if this interpretation is correct, the corresponding excited state is at 0·9 MeV. Similarly the peak at 5·6 divisions corresponds to an excited state at 2·0 MeV in one of the stable neon nuclei.

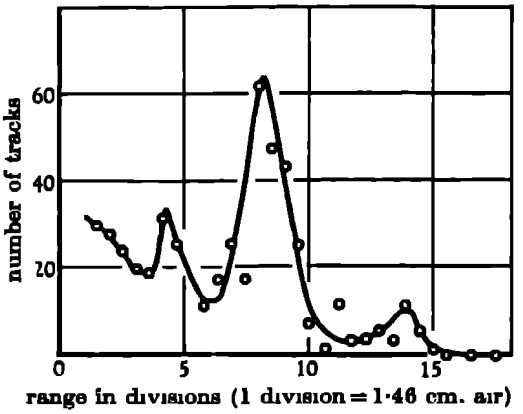


FIGURE 11. Inelastic scattering of protons from neon. Distribution in range of the inelastically scattered protons from observations at 50, 42 5 and 27·5°.

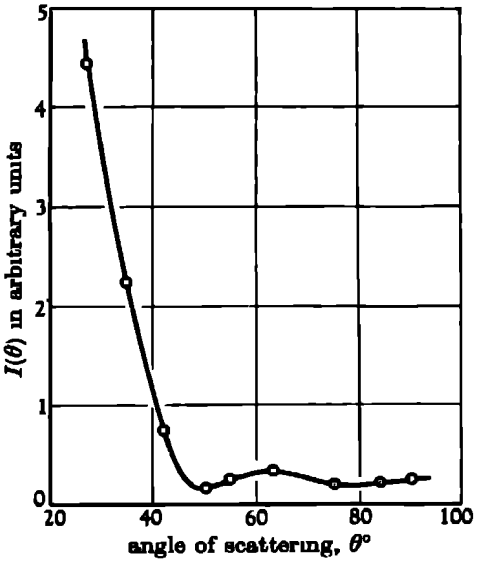


FIGURE 12. Inelastic scattering of 4 2 MeV protons by neon. Angular distribution of the intensity of the group corresponding to the formation of Ne^{20} in an excited state at 1·5 MeV.

We have examined the variation in intensity of the main group of inelastically scattered particles over the range from 40 to 90°, and the results are shown in figure 12. The sharp rise in the intensity at small angles emphasizes the importance of this angular region, a conclusion strongly reinforced by experiments on (d,p) reactions using fast deuterons In many scattering and disintegration processes with primary particles of high energy there appears to be a strong probability of the product particles being emitted at small angles to the direction of the primary beam

Observations of a similar character on inelastic scattering have been made by Wilkins & Kuerti (1940), Wilkins & Wrenshall (1940), Wilkins (1941) and Dicke & Marshall (1943). Wilkins studied the inelastic scattering of 6 MeV protons incident on a foil of magnesium, measuring the distribution in intensity of the inelastically scattered protons in the interval from 60 to 150° and showing that there is a maximum in the neighbourhood of 90° It seems clear that in experiments of this kind we are approaching phenomena of considerable complexity and that it is important to make a detailed examination of the variation in intensity of the different groups of inelastically scattered protons over the whole angular range, using primary protons of a number of different energies. The experiments with 6 MeV protons show that it is a considerable advantage to work at higher energies than were available in our experiments. At 4 MeV the ratio of inelastic to elastically scattered particles is inconveniently small with elements of higher atomic number than 10 Our experience suggests that it would be possible to use the present methods with primary proton energies up to 15 MeV and this should make it possible to make a detailed examination of the excited states of even heavy stable nuclei

7 Experiments with argon

An exposure was taken with the camera filled with cylinder argon at a pressure of 5 cm of Hg The range distribution of the tracks in this plate shows a sharp peak corresponding to the elastically scattered protons, and the variation of its intensity with angle follows the Rutherford law closely for angles of scattering between 20 and 70°, see I, p. 16. In addition there is a small number of tracks of shorter range, of which the number at any one angle is too small for a reliable analysis In figure 13 the distribution in range of all tracks of length less than 15 cm. standard air is shown irrespective of the angle of emission They form two distinct groups at 1·5 and 5·5 divisions respectively, corresponding to ranges in air of 3·0 and 8 6 cm respectively when allowance is made for the stopping power of the gas in the chamber, etc

The range of both these groups is greater than that to be expected for α-particles from the reaction

$$A^{40} + H^1 \rightarrow Cl^{37} + He^4$$

for which the release of energy calculated from the masses (Livingston & Bethe 1937) is 1·28 MeV. We therefore suggest that they are due to the inelastic scattering

of the incident protons the energy of the corresponding excited states in A^{40} being at 1 5 and 2 4 MeV respectively It is necessary to obtain results with more energetic protons before attempting a more detailed analysis

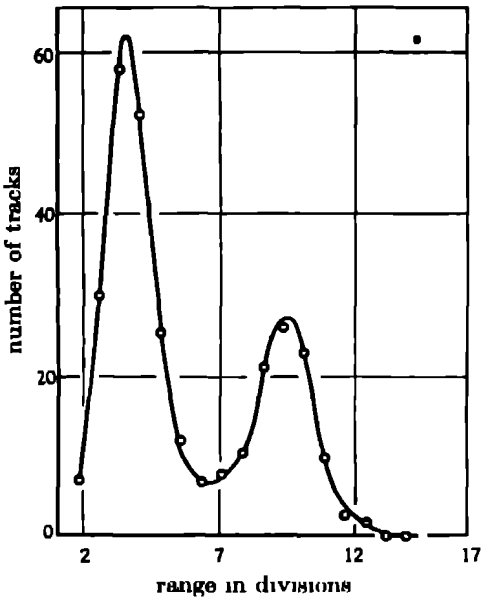


FIGURE 13. Short range particles from argon

8 SCATTERING BY CHLORINE AND BROMINE

The exposures for these experiments were made by filling the camera with the vapours of $CHCl_3$ and CH_3Br Because of the high nuclear charge of the chlorine and bromine atoms relative to those of the other elements in these compounds, most of the scattered protons arise from collisions with the former The results are thus very little affected by the presence of carbon and hydrogen nuclei in the scattering vapour. In plate P 16, obtained with chloroform as the scattering vapour, there are indications of the presence of short-range groups but their intensity is even lower relative to the elastically scattered protons than in the experiments with argon, and we have not measured a sufficient number of tracks to give a reliable range distribution

In the experiments with bromoform it was found that near $\theta = 90°$ the energy of the main proton group did not vary with angle in the way to be expected if it is due to the elastic scattering of the primary particles by bromine nuclei Thus the energy at the larger angles of scattering was lower than that expected by about 300 keV This was found to be due to the presence in the main group of an appreciable fraction of protons elastically scattered from carbon Because of the slow variation of the intensity of these protons with angle, due to the S-wave scattering, their number relative to those scattered from bromine becomes appreciable beyond 70° and,

although they are not resolved from the latter, they produce a displacement of the mean energy of the group

The experiments described in this paper were undertaken as part of an extended programme in collaboration with Sir James Chadwick and his colleagues in Liverpool, but, as a result of the war, we alone were able to proceed with the measurements. We are indebted to them for their agreement that we should publish the results here presented.

The measurements were made in the H. H. Wills Physical Laboratory, University of Bristol and one of us (A N M), is indebted to Professor A M. Tyndall for hospitality in the laboratory during the years 1940 to 1942.

References

Bonner 1940 *Proc Roy Soc* A, **174**, 339.
Buckingham & Massey 1941 *Proc Roy Soc* A, **179**, 123
Chadwick, May, Pickavance & Powell 1944 *Proc Roy Soc.* A, **183**, 1
Dicke & Marshall 1943 *Phys Rev.* **63**, 86.
Heydenburg & Ramsey 1941 *Phys Rev* **60**, 42.
Landau & Smorodinsky 1944 *J. Phys* **8**, 3, p 154.
Livingston & Bethe 1937 *Rev. Mod Phys.* **9**, 373
May & Powell 1947 *Proc Roy Soc* A, **190**, 170
Mott & Massey 1933 *The Theory of Atomic Collisions*, p 275 Oxford. Univ. Press
Powell 1940 *Nature*, **145**, 155.
Powell 1943 *Proc. Roy Soc* A, **181**, 344
Powell, May, Chadwick & Pickavance 1940 *Nature*, **145**, 893
Taschek 1942 *Phys. Rev* **61**, 13
Tuve, Heydenburg & Hafstad 1936 *Phys Rev.* **50**, 806.
Wilkins 1941 *Phys. Rev* **60**, 365.
Wilkins & Kuerti 1940 *Phys Rev* **57**, 1082.
Wilkins & Wrenshall 1940 *Phys Rev.* **58**, 758

The elastic scattering of 6·5 MeV deuterons by deuterium, helium and other light elements

By K. M Guggenheimer, H Heitler and C. F. Powell

The H. H. Wills Physical Laboratory, University of Bristol

(*Communicated by A M. Tyndall, F.R.S —Received 4 June* 1946)

The methods previously described have been applied to a study of the scattering of 6 5 MeV deuterons by deuterium, helium, carbon, nitrogen, oxygen and neon In the case of deuterium, helium and oxygen the elastically scattered deuterons can be clearly distinguished at all angles of scattering from groups of particles due to other nuclear processes. In experiments with these gases it is thus possible to determine the angular variation of the scattered intensity The results show that P and D terms make appreciable contributions to the amplitude of the scattered wave This result is in contrast with that obtained previously in experiments with protons of approximately the same speed, in which it was found that the scattered intensity could be described by an interaction of the Coulomb term with an S-wave only

In experiments with carbon, nitrogen and neon, the peaks in the range distributions due to the elastically scattered deuterons are confused with those produced by protons or α-particles from $(d\text{-}p)$ and $(d\text{-}\alpha)$ reactions, and the results are less reliable or are confined to restricted intervals in the angle of scattering. The Q values of the reactions giving rise to the proton groups have been determined and values thus obtained of the energy of a number of excited states in the product nuclei

A study of the variation with the angle of scattering of the range of deuterons elastically scattered by helium nuclei is shown to afford a method of investigating the range-energy relation for an emulsion which has advantages over that previously employed

1. Introduction

In previous papers (Chadwick, May, Pickavance & Powell 1944, May & Powell 1947 (see p 170), Heitler, May & Powell 1947, (see p 180)) which will be referred to as I, II and III respectively, experiments have been described on the scattering of the beam of 4 2 MeV protons delivered by the Liverpool cyclotron Gaseous 'targets' of the different elements were employed, the scattered particles being detected by the tracks which they produced in a photographic emulsion This paper describes similar observations with 6 5 MeV deuterons As in the case of the work with protons, the experiments constitute only a preliminary survey of the extensive field of investigation which it will be necessary to examine in detail to provide sufficient material for testing any theory of the various collision processes.

The details of the exposures given to the plates from which the present measurements were obtained are given in table 1. Since no provision had been made for measuring the current in the defined deuteron beam, the exposure given to any plate is defined in terms of the product of the time of exposure and the current in the main beam as measured by the collecting electrode, see paper I, figure 2 The current in the defined beam is of the order of 1 % of that recorded

The 23 exposures in series D were obtained in the course of 2 days in January 1941, most of the measurements on the resulting plates being made in 1941 and 1943 Attention may again be directed to the great saving in the operation-time of the cyclotron which the photographic method allows for experimental work of this type

TABLE 1

plate number	scattering gas	pressure (cm of Hg)	exposure (μA-min)
D.19	helium	50	15
D 23	deuterium	3	50
D 9	oxygen	30	10
D 10	oxygen	30	60
D 6	nitrogen	16 5	40
D 7	acetylene	30	10
D 8	acetylene	30	40
D 20	neon	7	50

In the work with protons the scattered particles at any angle, even in the most complicated cases, consisted of only a few homogeneous groups due to the elastic or inelastic collision of the primary particles by the target nuclei In the present experiments we meet with phenomena of much greater complexity, for the elastically scattered deuterons are in general accompanied by α-particles and protons from $(d\text{-}\alpha)$ and $(d\text{-}p)$ reactions. Further, these disintegration particles are not homogeneous but are distributed in a number of groups of different energy corresponding to the formation of the product nuclei in one or another of various possible states

As a result of the presence of the disintegration particles, the curves representing the distribution in range of the tracks observed at any angle of scattering show a number of peaks which are sometimes imperfectly resolved from one another. In experiments with deuterium, helium and oxygen, however, the elastically scattered deuterons can be clearly distinguished from other groups and their intensity, $I(\theta)$, measured in a wide range of angles. On the other hand, in the exposures with scattering gases containing carbon, nitrogen and neon, the interpretation of the results is made difficult by the presence of proton or α-particle groups of nearly the same range as that of the elastically scattered deuterons As a result the intensity measurements are less reliable or are confined to restricted intervals in the value of θ

2. ENERGY OF THE PRIMARY DEUTERONS

In I it was shown that the mean point of scattering of the primary particles changes with the angle of scattering and that there is thus a steady fall in the value of the quantity E_0^θ as θ increases from 10 to 170° In the previous work with protons the variation was very small but it is appreciable in the present experiments owing to the higher gas pressures employed in the camera The actual values of E_0^θ in a particular exposure can be determined in the following manner·

The mean range, $\rho_{\bar{\theta}}^\theta$, of the elastically scattered particles at a particular value of θ, $\bar{\theta}$, is determined by measuring the length of the tracks produced by the particles

in the appropriate area of the plate Allowance is made for the loss in range which the particles suffer, in traversing the foil and the gas in the camera, before reaching the plate From the observed value of ρ_s^θ, E_s^θ can be determined from the range-energy relation for the particles. The value of E_0^θ can then be found since it is simply related to E_s^θ, θ and the masses of the two particles interacting in the collision ρ_0^θ can then be deduced from the range-energy relation.

If r is the internal radius of the scattering tube, s the stopping power of the gas in the camera at the pressure employed during the exposure, then the value of $\rho_0^{\frac{1}{2}\pi}$, corresponding to $\theta = \frac{1}{2}\pi$, is given by the equation

$$\rho_0^\theta = \rho_0^{\frac{1}{2}\pi} + s \cdot r \cot\theta$$

The value for ρ_0^θ is employed to determine $\rho_0^{\frac{1}{2}\pi}$ and then ρ_0^θ can be determined for all other angles. Values of E_0^θ follow from the range-energy relation

In making such calculations, the value of s for any gas is determined from the values of the atomic stopping power of the atoms composing it given by Livingston & Bethe (1937). It is assumed, as a sufficient approximation, that the atomic stopping powers of the various elements do not vary with the speed of the particles in the range of energy involved

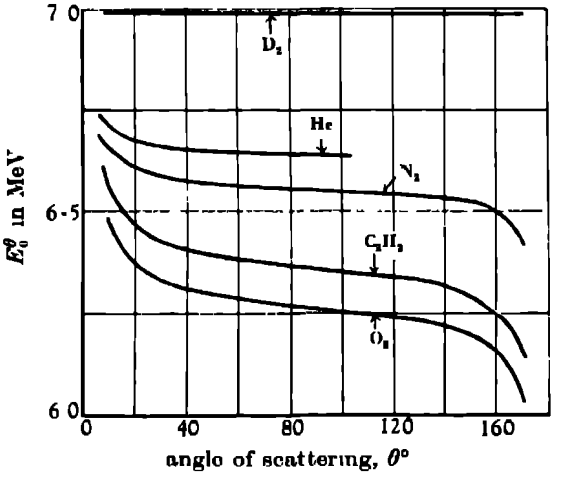


FIGURE 1 Variation with θ of E_0^θ, the mean energy of the primary deuterons at the mean point of scattering

The resulting values of E_0^θ for the exposures from which the present measurements were obtained are shown in figure 1 It will be seen that in the case of the exposure with oxygen, E_0^θ varies from 6 5 MeV at $\theta = 10°$ to 6 0 MeV at $\theta = 170°$ In the interval from $\theta = 40°$ to $\theta = 140°$, however, the change in the magnitude of E_0^θ is only 0 13 MeV, a value which we shall show to be much less than the spread in energy of the primary deuterons emerging through the defining slits from the cyclotron tank It is in this latter range of angles that the curves representing the

observed variation of the intensity of the elastically scattered deuterons display the most interesting features and we may conclude that the variation in E_0^θ will have no serious influence on the results

EXPERIMENTS WITH HELIUM

3 The range energy relation for the emulsion

We have examined the plate taken with helium as the scattering gas in considerable detail as it provides a means of testing certain important features of the method. The observations show that the elastically scattered deuterons give rise to the most prominent group in the range distributions at all angles of scattering and figure 2 shows a typical example for $\theta = 12\ 5°$ It is convenient to transform the observed distribution in range to a distribution in energy, assuming all the particles to be protons and using the appropriate range-energy relation The advantage of this procedure is that it gives a less extended scale, the peaks due to different groups all having approximately the same width at half maximum In such a distribution the position of the peak due to the elastically scattered deuterons will not of course indicate the true energy of the particles It will be seen from figure 2 that, in addition to the elastically scattered deuterons, tracks of shorter range are also present We attribute these to protons arising from the reactions represented by the equations

$$He_2^4 + H_1^2 \rightarrow He_2^3 + H_1^1 \quad \text{and} \quad He_2^4 + H_1^2 \rightarrow He_2^4 + H_1^1 + n_0^1 \quad \text{(see below)}.$$

The range of the α-particles produced by the impact of the primary deuterons on helium nuclei is too low for them to reach the photographic plate in the conditions of the exposure

Figure 3 shows the observed variation with θ of the mean range of the elastically scattered deuterons, ρ_s^θ Simple mechanical considerations (see Appendix) show that the energy of these particles, E_s^θ, should vary with θ according to the relation

$$E_s^\theta = \tfrac{4}{9} E_0^\theta \frac{\sin^2\phi}{\sin^2\theta} \quad (1), \quad \text{where} \quad \phi = \delta + \theta \ \text{and} \ \sin\delta = \tfrac{1}{2}\sin\theta.$$

In previous work it has been assumed that the range of a particle in the emulsion bears a constant ratio to its range in air irrespective of the nature of the particle and of its speed. Although it is clear that this assumption can only be approximately true, the previous experiments suggested that any errors to which it leads are insignificant in view of the limitations to the precision of the method in its present stage of development In accordance with this assumption the full line in figure 3 is plotted from values of the range of the scattered particles based on calculated values of E_s^θ, assuming that the stopping power of the emulsion is constant for deuterons of all energies and equal to the value for protons previously determined (10μ = 14·6 mm of standard air, Powell 1943). The values of the mean energy of the primary deuterons for different values of θ, E_0^θ were calculated with the same

assumption, from the observed range of the scattered particles at $\theta = 15°$. The experimental point at 15° therefore necessarily coincides with the curve.

The accuracy of the range determinations is clearly displayed in figure 3, the average departure of the experimental points from the curve being less than 2 mm air. Such an error corresponds to only 0 05 MeV in the estimate of the mean energy

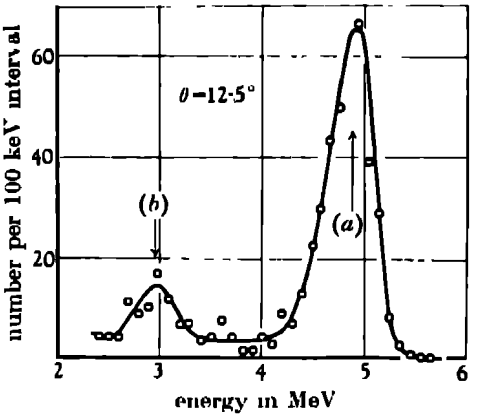


FIGURE 2 Observed range distribution at $\theta = 12\ 5°$ transformed to an energy scale assuming all particles to be protons Peak (*a*) is due to the elastically scattered deuterons, (*b*) to disintegration protons

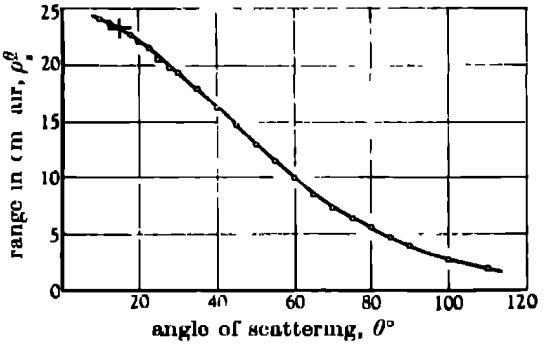


FIGURE 3 Variation of the mean range of the elastically scattered deuterons with angle of scattering, θ The full line is calculated from the observed range at 15° assuming the stopping of the emulsion to bear a constant ratio to that for air for deuterons of all energies in the interval from 2 to 7 MeV.

of the particles composing any group. Further, the consistency between the experimental points and the full line indicates that the assumption of a constant stopping power is justified, with the precision that the method at present affords, for deuterons with energies between 6·5 and 2 MeV The stopping power of a material for protons has the same value as that for deuterons of the same speed The observations therefore provide additional support for the assumption that the stopping power is

independent of the energy in the case of protons and deuterons. The previous experiments (Powell 1943) indicated that such an assumption is also true for α-particles

The constancy of the stopping power of the ordinary 'half-tone' emulsion is presumably due to the fact that most of the atoms are of carbon, nitrogen and oxygen, with a mean atomic number close to the value for air P Cuer, in this laboratory, has made calculations of the range-energy relation for the new 'concentrated half-tone' emulsions (Powell, Occhialini, Livesey & Chilton 1946) in which the proportion of silver halide is greatly increased He finds that very considerable variations of the stopping power are to be anticipated. For work of greater precision, which becomes possible with the new emulsions, it will be important to make an accurate determination of the range-energy relation for protons, α-particles, etc , in the new emulsions For this purpose the present method, which enables us to measure the lengths of the tracks of groups of deuterons of any desired mean energy in a continuous range of values, has certain advantages over that in which the observations are restricted to a few homogeneous groups from particular nuclear reactions Similar measurements with protons can be made by observations of the scattering of these particles by deuterium and helium

4 Inhomogeneity in the energy of the primary deuterons

In I, the angular resolving power of the method, which depends on the distribution of the actual angles of scattering suffered by the particles producing the tracks recorded in any small area of the plate, was determined A second important technical aspect of the method is its capacity to distinguish homogeneous groups of similar particles of different energy This question has also been examined in previous experiments (Powell 1943) by measuring the distribution in length of the tracks produced in the emulsion by homogeneous groups of fast protons and α-particles from various nuclear transformations Using the range-energy relation such observations can be transformed to a distribution in energy The width at half-maximum of the peak thus obtained, in experiments with a given homogeneous group, is defined as the 'energy resolving power' and the previous experiments indicated that this quantity has a value less than 0 3 MeV for proton groups of any energy in the interval from 2 to 13 MeV, when using the ordinary 'half-tone' emulsion

In the present experiments the width at half-maximum of the peak in the curve representing the distribution in energy of the deuterons scattered by helium at $\theta = 20°$ is 0·7 MeV, a value much greater than that obtained previously in the experiments with protons. It can be shown, however, that this result is not due to a reduction in the precision of the range measurements, but is caused by a lack of homogeneity in the energy of the primary deuterons. Thus figure 4 shows typical examples of the distribution in energy of the deuterons scattered from helium at various values of θ in the range from 15 to 100°.

Figure 4 shows that there is a steady decrease in the width of the peaks at half-maximum, ΔE_0^{θ}, as the angle of scattering increases. This effect is represented

graphically in figure 5 which shows the variation of ΔE_s^θ as a function of the mean energy of the particles in the corresponding group A decrease of this character is to be expected if there is a spread in energy of the primary particles. Thus, from equation (1), it follows that ΔE_s^θ is related to the corresponding quantity for the primary particles, ΔE_0^θ, by the equation

$$\Delta E_s^\theta = \Delta E_0^\theta \frac{E_s^\theta}{E_0^\theta} \tag{2}$$

We have estimated the values of the half-width of the peaks to be expected for different values of ΔE_0^θ and the line in figure 5 is calculated on the assumption that this quantity has a value 0 6 MeV, and that, in accord with the results of previous experiments, the half-width of the peak corresponding to a strictly homogeneous group of particles is 0·3 MeV The good agreement of the experimental points with this line shows that the present measurements of the lengths of the tracks and the energy-resolving power are of the same order of precision as that previously attained.

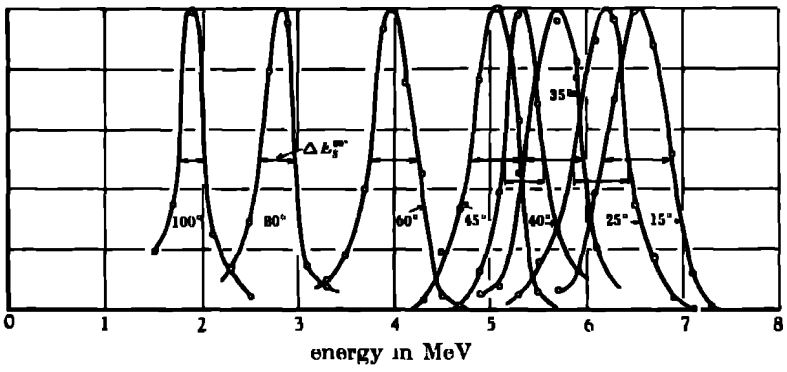


FIGURE 4 Energy distribution of elastically scattered deuterons at various angles of scattering, θ The peaks are scaled to give the same maximum intensity. They have been deduced from the observed range distributions, using the appropriate range-energy relation for deuterons and assuming that the emulsion has a constant stopping power ($10\mu = 1$ 46 cm air).

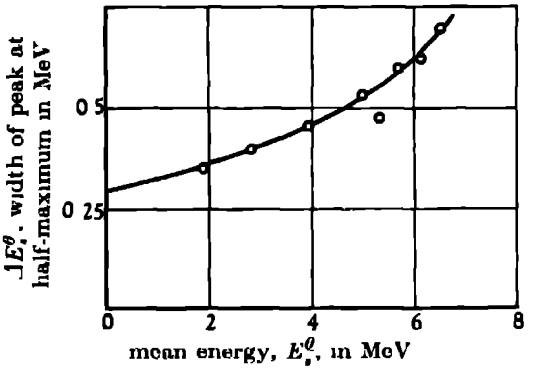


FIGURE 5 Diminution of observed width at half-maximum with improved homogeneity of scattered deuterons.

5 Angular distribution of the elastically scattered deuterons

The observed angular variation of the intensity of the elastically scattered deuterons, when transformed to the centre-of-mass co-ordinate system, is represented in figure 6 (*a*). The results are based on the measurement of a total of 5000 tracks at the various values of θ. In a previous paper it was shown that the scattering of 4·2 MeV protons by oxygen, nitrogen and carbon nuclei could be accounted for in terms of a formula of Mott & Massey which expresses the scattered intensity in terms of the interaction of the Coulomb term with S, P and other waves. In this formula the amplitude and phase of the S-wave are both defined by a single parameter ξ_0 and curves were given (III, figures 9, 10) corresponding to various values of ξ_0 from 0 to π, the values of the constant α being chosen to correspond to the collision of 4 2 MeV protons with the different nuclei.

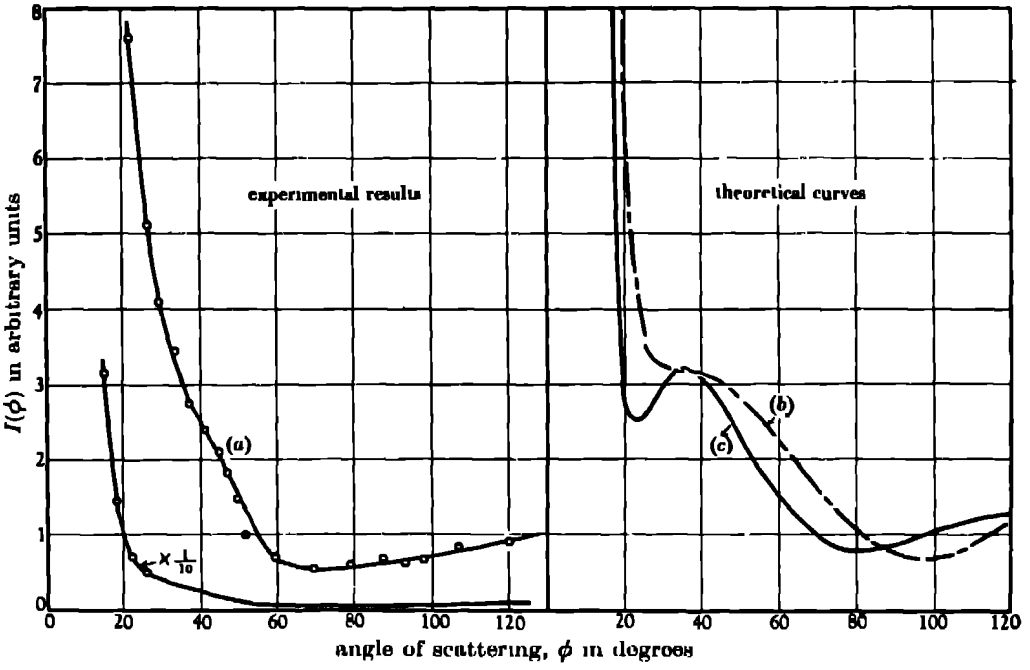


Figure 6 Elastic scattering of 6 6 MeV deuterons by helium. The theoretical curves are calculated from the formula of Mott and Massey, ξ_0, ξ_1 and ξ_2 being given the values indicated, $\alpha = 0\ 25$. Curve (*b*) $\xi_0 = \frac{3\pi}{4}$, $\xi_1 = \frac{\pi}{4}$, $\xi_2 = 0$. Curve (*c*) $\xi_0 = \frac{3\pi}{4}$, $\xi_1 = \frac{\pi}{4}$, $\xi_2 = \frac{\pi}{12}$.

In the case of the d-α collision, with 6 6 MeV deuterons, the appropriate value of α is 0·25 and we have calculated similar curves from the formula, putting the parameters ξ_1, ξ_2, etc , equal to zero. The resulting family of curves is similar to that for $\alpha = 0\ 63$ previously given and it is clear that the present experimental curve shows more complicated features than can be described by such a simple formula. We have

therefore made further calculations taking into account contributions from waves of higher angular momentum, and the results are included in figure 6, curves (*b*) and (*c*). The general form of the experimental curve is then found to be reproduced by the theoretical curves, and, in particular, the point of inflexion at $\phi = 30°$, the fall to a minimum at 70° and the subsequent rise Curve (*c*) is calculated with

$$\xi_0 = \frac{3\pi}{4}, \quad \xi_1 = \frac{\pi}{4} \quad \text{and} \quad \xi_2 = \frac{\pi}{12}$$

and shows that a small contribution from the *D* wave has an important effect on the calculated intensity distribution We have not attempted, at this stage, to determine values of ξ_0, ξ_1 and ξ_2 which reproduce most closely the detailed features of the experimental results. It is clear, however, that the latter can only be accounted for in terms of an interaction of the Coulomb term with *S*, *P* and *D* waves This result appears to be a general feature of the scattering of 0·5 MeV deuterons by the light elements in contrast with that obtained in the experiments with 4 2 MeV protons in which the particles have approximately the same speed

6. The dissociation of deuterons by impact with helions

We referred in a previous paragraph to the particles of shorter range which appear at small values of θ and which we attributed to disintegration protons The discriminating power of the emulsion, the difference in the density of tracks produced by protons, deuterons and α-particles of the same range, is not sufficiently great with the ordinary 'half-tone' emulsion to enable us to identify the short-range tracks by inspection and we have not attempted the method of grain counting for this purpose in this particular case If we assume that the particles are protons from the reaction $He^4 + H^2 \rightarrow He^5 + H^1 + Q$ we can transform their observed distribution in range at any particular angle, to an energy scale in terms of Q, the energy release in the reaction (see Appendix). The results for the distribution at different angles can then be assembled together and figure 7 shows the results of measurements of 400 tracks at 12 5, 15, 25, 35, 40 and 45° There is good evidence from this figure of a peak at $Q = -2\,9$ MeV.

If the group of protons composing the peak is attributed to the reaction $He^4(d, p)He^5$, the corresponding mass of He^5 is 5 0137 mass units, a value equal to that obtained by Williams, Shepherd & Haxby (1937) in experiments on the α-particles emitted from a lithium target under deuteron bombardment These authors determined the distribution in energy of the emitted α-particles using counter methods of detection They found evidence for a broad peak, superposed on the continuous distribution of particles from the reaction $Li^7 + H^2 \rightarrow 2He^4 + n^1$, which they attribute to α-particles from the alternative reaction

$$Li^7 + H^2 \rightarrow He^4 + He^5.$$

The mass of the He^5 nucleus was determined from the mean energy of the α-particles composing the peak The authors suggest that the breadth of the peak is to be

associated with the short mean life time of the recoiling He^5 nuclei before each dissociates into an α-particle and a neutron, a process which the observed mass indicates to be energetically possible Corresponding to the short lifetime, the mass of a He^5 nucleus may have any value in a range defined by the uncertainty principle, and the momentum and energy of the α-particle recoiling from it is correspondingly indefinite

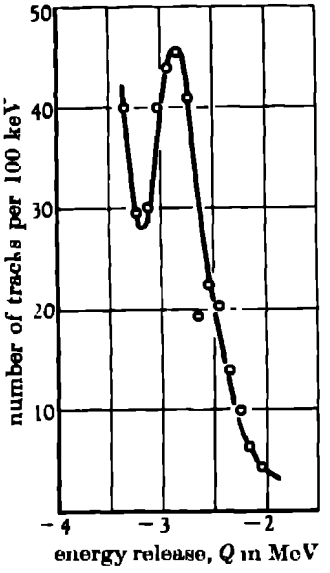


FIGURE 7 Distribution in energy, Q, of the protons produced by the collision of 6 6 MeV deuterons with helium nuclei

It is reasonable to suppose that the proton group shown by the curves in figure 7 can be interpreted in an analogous manner as due to the formation of short-lived He^5 nuclei The continuous energy distribution of protons on which the peak is superposed is, we suggest, to be attributed to the dissociation of the intermediate nucleus into three particles in accordance with the equation

$$He^4 + H^2 \rightarrow He^4 + H^1 + n^1.$$

Whether or not the intermediate formation of He^5 plays a role in the collision, the final result is the production of a neutron and a proton so that the process can be regarded as one in which the deuteron is dissociated by impact with a helion

EXPERIMENTS WITH DEUTERIUM

7. DEUTERON-DEUTERON SCATTERING

The scattering of 7·0 MeV deuterons by deuterium has been studied over the angular range from 10 to 60°, the observed variation of $I(\theta)$ with θ being shown in figure 8. The peak due to the elastically scattered deuterons forms the dominant

feature of the range distribution at all angles, the number of protons and tritons from the reaction $H^2 + H^2 \rightarrow H^3 + H^1$ being relatively very small. It is a satisfactory feature of the measurements that the $I(\phi)$, ϕ curve, in which the results are transformed to the centre of mass co-ordinate system, is symmetrical about the line $\phi = \frac{1}{2}\pi$ within the accuracy of the measurements, corresponding to the equivalence of the two interacting particles.

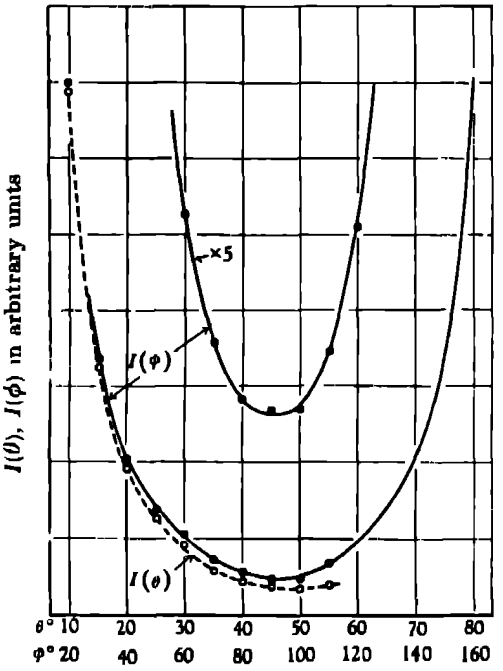


FIGURE 8. The collision of 7 MeV deuterons with deuterium. Angular distribution of the elastically scattered deuterons

A comparison of the results with those for the proton-proton collision (see II, p 170) shows that we cannot account for the scattering in terms of an interaction between the Coulomb term and an S-wave alone, and that contributions from waves of higher angular momentum are of importance This is to be expected since the de Broglie wave-length for the collision is $2\,4 \times 10^{-13}$ cm. whilst the estimated range of the nuclear forces is about 7×10^{-13} cm. It is therefore reasonable to expect strong scattering for the wave of angular momentum $l = 2$ (Livingston & Bethe 1937) Because of the symmetry of the two particles, only contributions from spherical harmonics of even order will appear, but we have not yet attempted to determine their relative importance Since the final expression for the scattered intensity appears as the square of the sum of a number of terms involving the different harmonics, it is not possible to deduce the magnitude of these terms directly from the observations and a method of trial and error has to be adopted involving laborious computations

For comparison with theoretical results the actual values of the scattered intensity at the various values of ϕ are shown in table 2

TABLE 2. DEUTERON-DEUTERON SCATTERING AT 6 6 MeV

$\theta°$	$\phi°$	$I(\phi)$	$\theta°$	$\phi°$	$I(\phi)$
10	20	42 0	40	80	3 69
15	30	20·8	45	90	3·54
20	40	13 2	50	100	3 54
25	50	9 01	56	112	4 54
30	60	6 93			
35	70	4 73			

$I(\phi)$ is in arbitrary units.

EXPERIMENTS WITH OXYGEN

8 RANGE DISTRIBUTIONS FROM THE d-O^{16} COLLISION, IDENTIFICATION OF THE DISINTEGRATION PARTICLES

Typical range distributions obtained from measurements on the plate taken with oxygen as the scattering gas are shown in figure 9, the angles of scattering being 30, 90 and 150°, respectively There are four distinct peaks labelled α, d, p_1 and p_2. The peak (d), which is clearly resolved at all angles, has a range which accords at all values of θ with that expected for elastically scattered deuterons We attribute the two peaks p_1 and p_2 to protons from the reaction O^{16} (d, p) O^{17}, the two groups corresponding to the formation of O^{17} in the ground state or in an excited state at 0 83 MeV (Cockcroft & Lewis 1936)

Confirmation of this interpretation of the origin of the long-range particles is provided by the observed changes in the ranges of the two groups with θ Owing to the relatively small mass of the oxygen nucleus, the motion of the centre of mass in the collision is appreciable and the mean range of the protons in the fastest group decreases from 85 cm of air at 10° to 47 cm at 170°. If, as in the case of the helium results, we transform the observations to the centre-of-mass co-ordinate system and express the proton distributions on an energy scale in terms of the energy release, Q, in the reaction, we obtain the result shown in figure 10. The curve is based on the measurements of a total of 4600 proton tracks, made up of 23 sets of 200 each, from observations at different angles ranging from 11 to 165°.

This method of analyzing the results provides an exceptionally severe test of the precision of the method Any departures from the basic assumption about the range-energy relation for the emulsion, and any inaccuracies in the determination of E_0^p, or of the measured length of the tracks, will lead to a loss of detail in the proton distribution In fact the width at half-maximum of the two peaks into which the experimental curve shown in figure 10 is resolved is 0 6 MeV, the value to be anticipated owing to the observed inhomogeneity in energy of the primary deuteron beam.

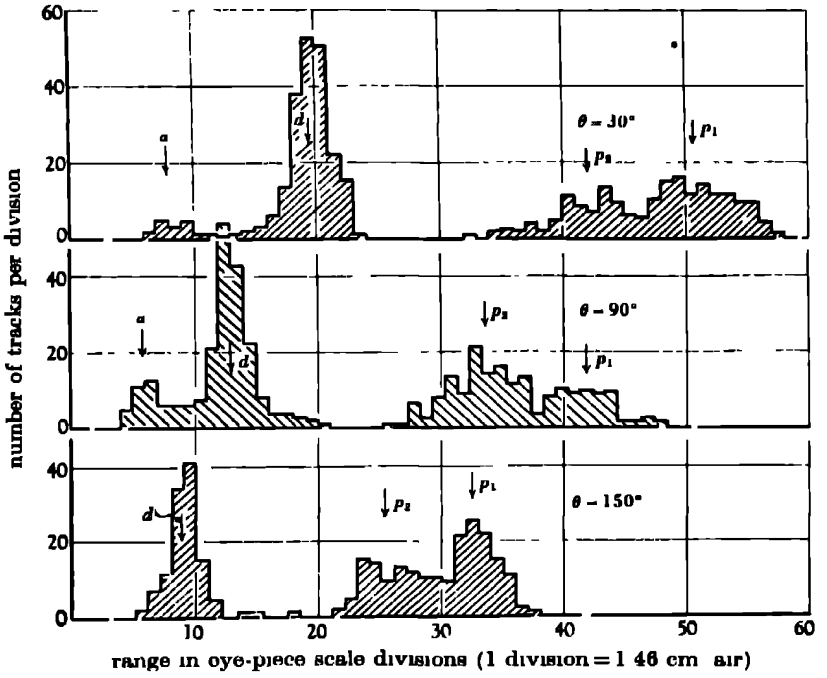


FIGURE 9 Distribution in range of tracks from the bombardment of oxygen by 6 3 MeV deuterons at $\theta = 30$, 90 and 150°.

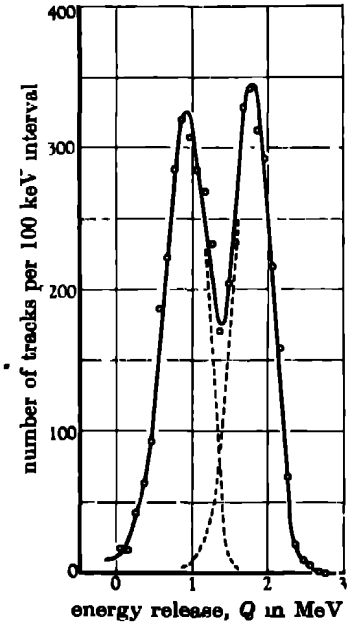


FIGURE 10. Distribution in Q, energy release in MeV, of protons from the reaction $O^{16}(d, p)O^{17}$. Results of measurement of 4700 tracks at 23 different angles in the interval from 11 to 165°.

The result shows that the particles are correctly identified as due to protons from the reaction O^{16} (d, p) O^{17} and gives a value of 0·85 MeV for the excited state in O^{17} corresponding to the peak of lower energy. This is in excellent agreement with the value, 0·83 MeV, obtained by Cockcroft & Lewis (1936) in experiments with slow deuterons and by Burcham & Smith (1938) from the reaction F^{19} (d, α) O^{17} The observed Q value for the formation of O^{17} in the ground state, 1·8 MeV, is a little less than that calculated from the best determinations of the masses, 1·95

We have not succeeded in establishing the nature of the particles of short range. If we regard them as protons from the reaction O^{16} (d, p) O^{17}, the energy of the corresponding state in O^{17} is not in agreement with any of the values found by Burcham & Smith (1938). The width of the peak in the range distribution observed at any angle is larger than that expected for a homogeneous group of particles, and whilst some of the tracks have a length equal to that expected for α-particles arising from the reaction O^{16} (d, α) N^{14} others are too long It is possible therefore that we are dealing with different groups of protons and α-particles The discriminating power of the new emulsions now available will enable us to decide this point in future experiments

9. Angular distribution of the deuterons elastically scattered from oxygen

The peaks in the range distributions produced by the elastically scattered deuterons are clearly resolved from other groups at all values of θ The results of intensity measurements are represented in figure 11 which shows them transformed to the centre-of-mass co-ordinate system It is clear that the form of the curves, in this case also, can only be interpreted in terms of S, P, and possibly other waves, interacting with the Coulomb term, and the dotted curve included in the figure is calculated from the Mott and Massey formula. The chosen value of the constant α, 0·68, is appropriate to the mean energy, 6·3 MeV, of the primary deuterons in collision with O^{16} The scales of the theoretical curves are adjusted so that they coincide with the experimental value at 15°.

In this collision process also, the main features of the experimental results are closely similar to those obtained from the theoretical formula when ξ_0, ξ_1 are given suitable values. The agreement is not exact but in view of the considerable changes in the form of the curves which can be brought about by giving ξ_2 a small value different from zero, the agreement can be considered to be satisfactory The wavelength associated with the primary deuterons with a mean energy of 6·3 MeV, in collision with oxygen, is $1\,4 \times 10^{-13}$ cm , and the range of the nuclear forces of the order of 6×10^{-13} cm. It is therefore possible that waves corresponding to $l = 4$ make appreciable contributions to the scattered intensity. If this is so we should expect to find a rather rapid change in the form of the angular distribution with changing values of the energy of the primary deuterons. The angular variation of the intensity of the disintegration protons, which has previously been referred to as

displaying features of considerable complexity, forms the subject of a later investigation.

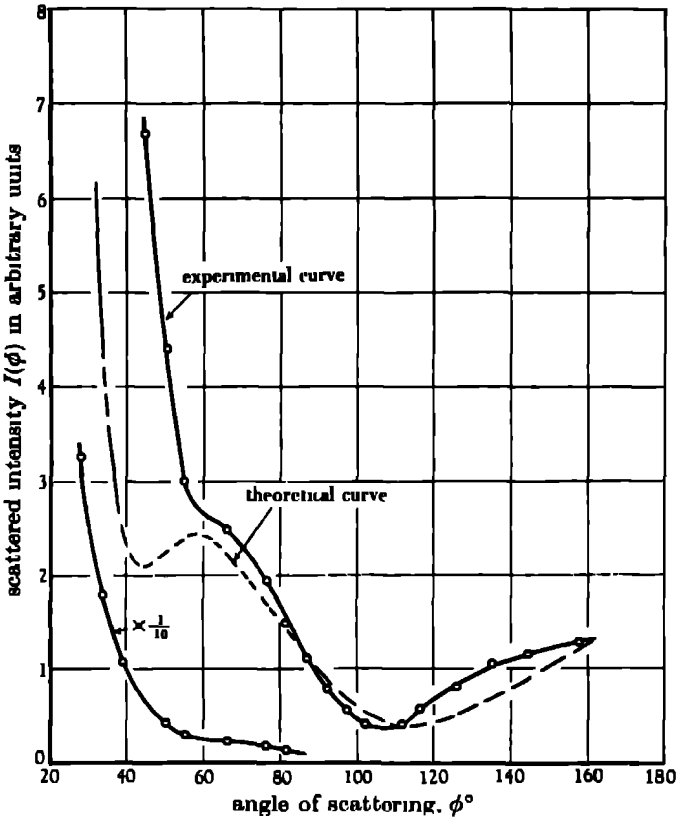


FIGURE 11 Elastic scattering of 6 3 MeV deuterons by oxygen.

Theoretical curve $\xi_0=\frac{3\pi}{4}$, $\xi_1=\frac{\pi}{6}$, $\alpha=0\ 68$

EXPERIMENTS WITH NITROGEN

10. THE RANGE DISTRIBUTION OF PARTICLES FROM THE d-N^{14} COLLISION; DISINTEGRATION PROTONS

The range distribution of the particles produced at any value of θ with nitrogen as the scattering gas is exceptionally complex Figure 12 shows three examples at $\theta = 105$, 120 and 135°, plotted for convenience on an energy scale assuming all the particles to be protons As in the work with oxygen we transform the results to the centre-of-mass system and again plot the results for the long tracks on an energy scale, assuming them to be produced by protons. We can compare the results from observations at different angles if we express them in terms of Q, the energy released in the reaction $N^{14}+H^2\rightarrow N^{15}+H^1+Q$, and figure 13 shows the resulting

distribution based on 3600 long tracks measured at various angles in the interval from 25 to 135°.

At small angles of scattering the dominant peak in the range distribution is that due to the elastically scattered deuterons, peak *d*, figure 12, and we can calculate the expected range of such particles at other angles The results represented in figure 13 refer only to measurements with particles of greater range than the values

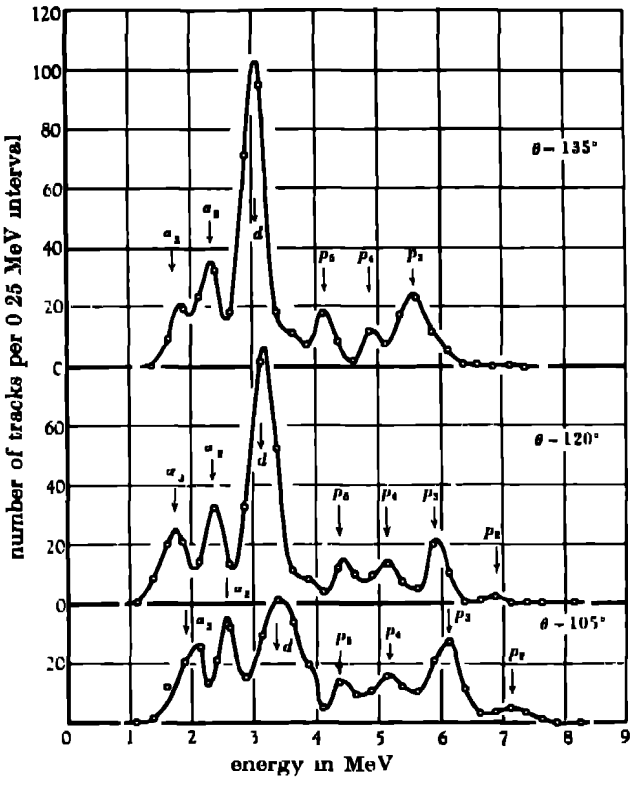


FIGURE 12. Results of measurements of the length of tracks from the *d*, N^{14} collision, transformed from a range distribution to an energy scale assuming all tracks to be produced by protons The peaks p_3, p_2, etc , are due to protons from the reaction N^{14} (d, p) N^{15}, α_2, α_3 are due to α-particles possibly from the reaction N^{14} (d, α) C^{12}, the C^{12} nucleus being left in an excited state The peaks *d* are largely due to elastically scattered deuterons.

so obtained It will be seen that there is a group p_0, of low intensity, at $Q = 8\ 55$ MeV, in good agreement with the known mass of N^{15} in the ground state Whilst the intensity of this group is certainly very low it is stronger, relative to the others, than is indicated in figure 13. The corresponding tracks are so long that with the angle of approach employed in the camera, a large proportion of them pass through the emulsion into the glass, and, since the range of such tracks is indefinite, they are not included in the analysis.

The prominent groups at $Q = 0\,3$, 1·3 and 3·5 MeV, marked p_4, p_3 and p_1 in figure 13, correspond to excited states in N^{15} at 8·2, 7·2 and 5·0 MeV respectively From the evidence provided by the variation with angle of the intensity of these groups, it seems likely that the first two are doublets which are unresolved in the present measurements because of the energy spread of the primary deuteron beam. In addition to these groups there is some evidence for a group at $Q = 2{\cdot}5$ MeV corresponding to an excited state in N^{15} at 6 0 MeV. The complexity of the proton distribution emphasizes the importance of making experiments with a more homogeneous deuteron beam so that the full resolution possible with the method can be employed. Especially with the new emulsions now available it should be possible

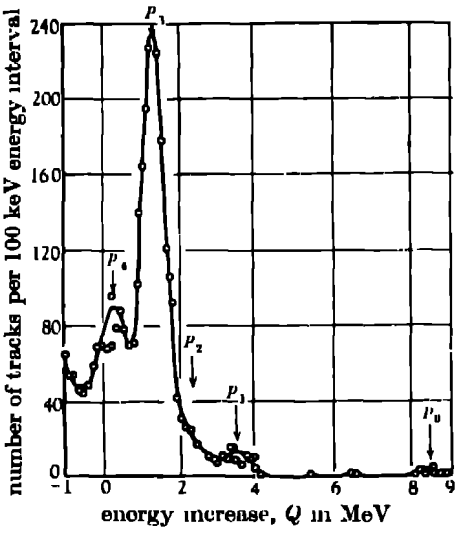


FIGURE 13 Protons from the reaction N^{14} (d, p) N^{15}. Distribution in terms of the energy release, Q 3600 tracks taken from observations at twelve different angles between 15 and 135°

to obtain peaks with a width at half-maximum of 100 keV which would allow us to resolve doublets with a separation greater than 0 2 MeV and to study the angular variation of their intensity, if sufficiently homogeneous beams of primary particles can be produced.

11. THE PARTICLES OF SHORT RANGE

The tracks of short range shown in figure 12 are distributed in a number of groups which correspond in range with that expected if they arise from the d, α reaction $N^{14} + H^2 \rightarrow C^{12} + He^4$. The α-particle group from this reaction, corresponding to the formation of the C^{12} in the ground state, should nearly coincide in range with the elastically scattered deuterons and the two groups would not be resolved in the conditions of the experiment. The excited states of C^{12} at 4·3 and 7·3 MeV are well

established (Bonner & Brubaker 1936; Powell 1943) and the ranges of the peaks marked α_1 and α_2 correspond to the formation of the C^{12} in these two states respectively. It is possible that there is also an unresolved, overlapping distribution of protons from the d, p reaction in this region, but grain counts on the tracks indicate that they are probably few in number relative to the α-particles.

12. Angular distribution of the elastically scattered deuterons

In spite of the possibility of confusion due to an overlapping α-particle group, we have studied the variation with angle of the prominent group which is mainly due to elastically scattered deuterons The results are shown in figure 14 and will be seen to display the same features as shown by the corresponding curve for oxygen. The

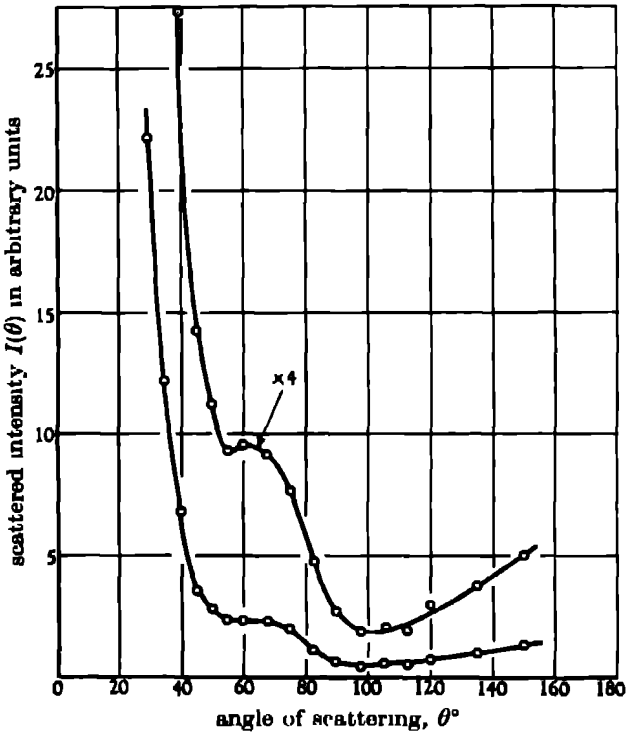


Figure 14 Angular distribution of 6·6 MeV deuterons scattered from N^{14}.

similarity is so close that if the two curves, normalized at 60°, are superposed they are found to coincide to within the accuracy corresponding to the statistical fluctuations arising from the finite number of tracks measured at any angle, a result which indicates that the disturbance due to α-particles, in the experiments with nitrogen, is small This striking result strongly suggests that the angular distribution

of the elastically scattered deuterons is not affected by competition from alternative processes and that it can be simply interpreted in terms of the wave-length appropriate to the collision and the range of the nuclear forces

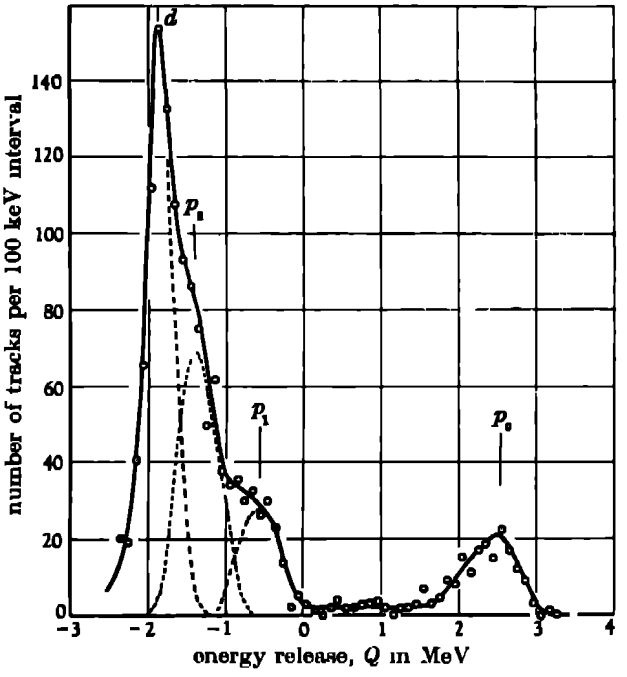


FIGURE 15 Energy distribution in Q of particles from the bombardment of carbon by deuterons. Results based on the measurement of the length of 1440 tracks from observations at 70, 87 5, 90, 105 and 120°, and transformed to an energy scale assuming all tracks to be produced by protons The peak d, due to elastically scattered deuterons should, from the known spread in energy of the primary beam, have a width of 0 4 MeV at half-maximum The proton peaks, p_1 and p_2, correspond to the formation of C^{13} in excited states at 3 15 and 3 95 MeV respectively

EXPERIMENTS WITH CARBON

13 THE RANGE DISTRIBUTIONS, THE DISINTEGRATION PROTONS

The experiments on the scattering from carbon were made with the camera filled with acetylene and as a result, at small values of θ, the observations are complicated by the presence of deuterons elastically scattered from hydrogen and protons projected in d-p collisions Above 30°, however, the presence of hydrogen gives rise only to projected protons and their range rapidly falls as θ increases

The deuterons elastically scattered from the carbon nuclei are accompanied by large numbers of disintegration protons and figure 15 shows the results, transformed from a range distribution to an energy scale, from observations on the length of the tracks at values of θ near 90°. The results are again shown in terms of the Q value, assuming all the particles to be protons from the reaction C^{12} (d,p) C^{13}. In these

circumstances the elastically scattered deuterons should appear as a peak with a half-width of about 0·4 MeV The complicated distribution of particles extending from $Q = -2$ MeV to $Q = 0$ can be analyzed into three components of which the most prominent, d, corresponds in energy to that expected for the elastically scattered deuterons The proton group of greatest range has a mean Q value of 2·6 MeV which agrees to within 0·1 MeV with the value calculated from the masses. The other peaks marked p_1, p_2, with Q values of -0 55 and -1·35 MeV, correspond to excited states in C^{13} of 3·15 and 3 95 MeV respectively Several experimenters have obtained evidence for an excited state in C^{13} at about 3 MeV (Chadwick, Constable & Pollard 1931; Feather 1932, Bennett, Bonner, Hudspeth, Richards & Watt 1941), but the state at 3·95 does not appear to have been reported previously

With carbon the study of the elastic scattering is rendered very difficult owing to the overlapping proton group. In view of the additional complications at small angles arising from elastic scattering by hydrogen, we have not attempted to study the variation with angle of the scattered intensity. Such an analysis may become possible with the new emulsions, using a more homogeneous primary beam, and a thin carbon target

EXPERIMENTS WITH NEON

14 In view of the great complexity of the range distributions obtained with neon as the scattering gas and the complications associated with the presence of more than one isotope, we have made no extended measurements on plate D 20 A preliminary survey shows that protons and α-particles are present in the scattered beam in considerable intensity and that it will be necessary to make experiments with separated isotopes

Preliminary experiments with argon and with a thin aluminium foil indicate that with these elements the elastically scattered deuterons give rise to the most prominent groups in the range distributions at all angles and that there is no serious confusion arising from the presence of disintegration particles

The experiments described in this paper were undertaken as part of an extended programme in association with Sir James Chadwick and his collaborators in Liverpool, and with Dr A. N. May Owing to the war we alone were able to proceed with the measurements and we are indebted to them for their concurrence in our proposal to publish the results here presented.

APPENDIX

The following approach to the problem of transforming the observations to the centre-of-mass co-ordinate system is well known. In the conditions of our experiments, however, exact numerical calculations are arduous because of the variation of E_0^c with θ. For the convenience of other observers we indicate the approximations which can be employed without a significant loss of precision

(*a*) *Elastic scattering*

We consider the case of the elastic scattering of deuterons by helium. Let v be the velocity of the primary deuterons measured in such units that the energy of the deuteron in MeV $= v^2$. In the centre-of-mass co-ordinate system the velocities of the deuteron and the α-particle are $\frac{2v}{3}$ and $\frac{v}{3}$, that of the centre of mass being $v/3$ relative to the laboratory. The velocity diagram in figure 16 (*a*) then shows that an angle of scattering ϕ in the centre-of-mass system corresponds to an angle of scattering θ of the deuteron and an angle of projection η of the α-particle, in the laboratory system where $\phi = \theta + \delta$

$$\frac{OA}{AB} = \frac{\sin\phi}{\sin\theta} \quad \text{and} \quad \frac{\sin\theta}{\sin\delta} = 2$$

It follows that

$$OA = \frac{2v\sin\phi}{3\sin\theta}$$

and the energy of the scattered deuterons at θ,

$$E_s^\theta = OA^2 = \tfrac{4}{9}E_0^\theta\frac{\sin^2\phi}{\sin^2\theta}.$$

Scattered intensity $I(\phi)$ in terms of $I(\theta)$

$$I(\theta)\sin\theta d\theta = I(\phi)\sin\phi d\phi,$$

$$I(\phi) = I(\theta)\frac{\sin\theta}{\sin\phi}\frac{d\theta}{d\phi}, \quad \text{but} \quad \frac{d\phi}{d\theta} = 1 + \frac{d\delta}{d\theta} = \frac{2\cos\delta + \cos\theta}{2\cos\delta},$$

$$I(\phi) = I(\theta)\frac{\sin\theta}{\sin\phi}\frac{2\cos\delta}{2\cos\delta + \cos\theta}.$$

(*b*) *Disintegration particles*

Relation between θ and ϕ Let the velocity of a primary deuteron be v where $v^2 = E_0$ Then in collision with an O^{16} nucleus its velocity in the centre-of-mass system is $\frac{8v}{9}$ and that of the oxygen nucleus $\frac{v}{9}$. The energy of the centre of mass is thus $\frac{1}{2}\ 18.\left|\frac{v}{9}\right|^2 = \frac{E_0}{9}$ and the energy available to the interacting particles $\frac{8E_0}{9}$

Suppose that as a result of the interaction a disintegration takes place with the emission of a proton with energy release Q MeV. In the centre-of-mass system let the proton velocity be v_1, and of the O^{17} nucleus $\frac{v_1}{17}$, then

$$\tfrac{1}{2}v_1^2 + \tfrac{1}{2}.17.\left(\frac{v_1}{17}\right)^2 = \frac{8E_0}{9} + Q,$$

or

$$v_1^2 = \frac{17}{9}\left(\frac{8E_0}{9} + Q\right) \tag{1}$$

Corresponding to given values of E_0 and Q, v_1 can thus be determined The relation between the angle of ejection of the proton in the centre-of-mass and laboratory systems of co-ordinates can then be found Thus from figure 16 (*b*) we see that

$$\phi = \theta + \delta, \tag{2}$$

and

$$\frac{\sin\delta}{v/9} = \frac{\sin\theta}{v_1}, \quad \sin\delta = \frac{v\sin\theta}{9v_1}, \tag{3}$$

so that the value of ϕ corresponding to any value of θ can be determined.

Determination of $I(\phi)$ from $I(\theta)$

$$I(\phi) = I(\theta)\frac{\sin\theta\, d\theta}{\sin\phi\, d\phi}.$$

From (2) $\quad \dfrac{d\phi}{d\theta} = 1 + \dfrac{d\delta}{d\theta}$, from (3) $\cos\delta\, d\delta = \dfrac{v}{9v_1}\cos\theta\, d\theta$.

Putting $\cos\delta = 1$, $\quad \dfrac{d\delta}{d\theta} = \dfrac{v\cos\theta}{9v_1}$, $\quad \dfrac{d\phi}{d\theta} = 1 + k\cos\theta$,

where $k = \dfrac{v}{9v_1}$, hence $\quad I(\phi) = I(\theta)\dfrac{\sin\theta}{\sin\phi}\dfrac{1}{1+k\cos\theta}$.

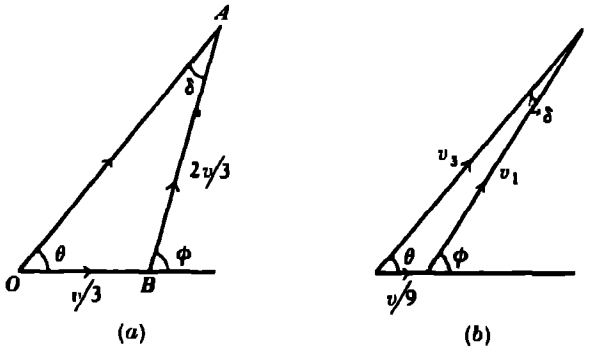


FIGURE 16

In the conditions of our experiment E_0^θ, and therefore v_1 and k, vary with θ. It is a sufficient approximation, however, to take the mean value $E_0^{\bar{\theta}}$, and the corresponding value of v_1, to determine a value of k which is treated as independent of θ.

Transformation of range distribution to an energy distribution in terms of Q

To transform a given range distribution of the disintegration protons to an energy scale in terms of Q we calculate the energy of the protons $E_p^{\theta,Q}$ emitted at angle θ and corresponding to an energy release Q. This energy $\frac{1}{2}v_3^2 = \frac{1}{2}v_1^2\dfrac{\sin^2\phi}{\sin^2\theta}$ (see figure 16 (*b*)), or, from (1)

$$E_p^{\theta,Q} = \frac{17}{18}\left(\frac{8E_0^\theta}{9} + Q\right)\frac{\sin^2\phi}{\sin^2\theta}.$$

At any value of θ, for positive values of Q, $E_p^{\theta,Q}$ varies nearly linearly with Q. It is therefore sufficient to calculate the proton energy for a number of values of Q, 0, 1, 2, 3, 4, etc. MeV for every 10° in the value of θ and to construct a family of curves showing $E_p^{\theta,Q}$ as a function of Q for these angles. Values for intermediate angles and energies can then be found by interpolation From the observed range distribution at a given angle the number of tracks in successive intervals corresponding to 0 1 MeV in the value of Q can then be determined.

For negative values of Q the values of v_1 become small and in this case it may be necessary to calculate values of $E_p^{\theta,Q}$ at closer intervals, say for every 0 3 MeV.

Similar methods can be employed with the d-α disintegrations.

References

Bennett, Bonner, Hudspeth, Richards & Watt 1941 *Phys Rev.* **59**, 781.
Bonner & Brubaker 1936 *Phys. Rev* **50**, 308.
Burcham & Smith 1938 *Proc. Roy Soc* A, **168**, 176.
Chadwick, Constable & Pollard 1931 *Proc Roy Soc.* A, **130**, 463
Chadwick, May, Pickavance & Powell 1944 *Proc Roy Soc.* A, **183**, 1
Cockcroft & Lewis 1936 *Proc. Roy Soc* A, **154**, 261.
Feather 1932 *Nature*, **130**, 237.
Heitler, May & Powell 1947 *Proc Roy Soc.* A, **190**, 180
Livingston & Bethe 1937 *Rev Mod Phys* **9**, 373.
May & Powell 1947 *Proc Roy Soc.* A, **190**, 170
Powell 1943 *Proc Roy. Soc.* A, **181**, 344.
Powell, Occhialini, Livesey & Chilton 1946 *J Sci. Instrum.* **23**, 102
Williams, Shepherd & Haxby 1937 *Phys Rev* **51**, 888

Ozonosphere temperatures under radiation equilibrium

By E. H. Gowan, *University of Alberta, Edmonton*

(*Communicated by G. M. B. Dobson, F.R.S.—Received* 2 *July* 1946
—*Read* 21 *November* 1946)

In two previous papers (Gowan 1928, 1930) calculating the effect of ozone on the temperature of the upper atmosphere, the height of the centre of gravity of the ozone layer was assumed to be about 45 km. Later, more reliable measurements were made (Gotz, Meetham & Dobson 1934) lowering this figure to 22 km, and also supplying the vertical distribution up to a little beyond 45 km. The results of the necessary revision of calculations were given to the Oxford Conference on Atmospheric Ozone (Gowan 1936). Strong (1939) and Summerfield (1941) have given a revision of the infra-red absorption of ozone and have taken the pressure into account. Elsasser (1942) and Dobson (1942) have discussed the absorption of water vapour in the atmosphere and Elsasser has justified the use of the square-root law for the pressure effect throughout the atmosphere for this vapour in the infra-red. The present paper gives the results of a further revised calculation making use of the new data, for latitude 50° N.

Method

The foundation of the method is the primary assumption that radiation equilibrium is closely approached about midday in the stratosphere as a whole, and that what convection exists is too slow to prevent the attainment of this equilibrium but is nevertheless sufficient to keep the composition substantially constant. The following equation represents this equilibrium in the form most convenient for computation

$$\int_0^\infty K'_\lambda S_\lambda d\lambda + \int_0^\infty K_\lambda E_\lambda d\lambda + \int_0^\infty K_\lambda A_\lambda d\lambda = 2\int_0^\infty K_\lambda B_{\lambda T_0} d\lambda$$

This equation applied to any horizontal layer where, for wave-length λ,
K'_λ is the fractional absorption for radiation in a parallel beam,
K_λ is the same for diffuse radiation,
S_λ is the solar radiation reaching the layer,
E_λ is the earth and troposphere radiation reaching the layer,
A_λ is the radiation to the layer from the rest of the stratosphere,
$B_{\lambda T_0}$ is the black-body radiation at T_0° K, which temperature has to be assumed for each layer at the beginning of the computation.

Graphical integration is necessary because of the irregular variation of K'_λ and K_λ. Successive approximation is necessary because A_λ depends on the temperature of

the entire ozonosphere except the layer being considered. For given latitude, season, amount of ozone, and amount of water vapour only the term containing A_λ needs recalculation at each stage of the successive approximation process.

Nine layers have been chosen from 11 to 15 km, then every 5 up to 55 km. The distribution of water vapour, carbon dioxide, and ozone in the nine layers was calculated on the basis of a preliminary estimate of temperature for average conditions at latitude 50° N These were then considered as 'optical layers', i e as containing this constant amount of gas Only small changes in height resulted from the final temperatures being different from the first estimate These have been neglected in the region considered

Data on quantities

In table 1 are shown the amounts of absorbing material in the nine layers. For oxygen and carbon dioxide the units of thickness are centimetres, in all the other columns they are microns For ozone and oxygen and carbon dioxide the amounts refer to normal temperature and pressure, for water vapour they are the equivalent thickness of liquid (compare Brunt & Kapur (1938) and Jaumotte (1936))

Table 1

height (km)	pressure at lower boundary (mm Hg)	ozone total 0·200 cm. (μ)	ozone total 0·280 cm. (μ)	O_3 (cm.)	water vapour 10 % (μ)	water vapour 40 % (μ)	CO_2 (cm.)
50–55	0 67	5	5	65 6	0 050	0·192	0 060
45–50	1·19	16	16	118	0 087	0 351	0 107
40–45	2·10	59	59	234	0 155	0 619	0·187
35–40	4·05	185	223	280	0 330	1 32	0·402
30–35	8 80	400	455	751	0·682	2 73	0·980
25–30	19·1	425	505	1480	1 92	7 68	2·12
20–25	41 6	355	470	4120	3·75	15 1	4 65
15–20	90·3	265	420	9730	8 30	33 2	10 10
11–15	168	156	304	16700	13·2	52 8	15·0

The water vapour was calculated as proportional to the total pressure (i.e completely mixed in the stratosphere) starting with 10 and 40 % relative humidity at the tropopause. The ozone was computed from the curves given by Gotz *et al* (1934). Carbon dioxide was calculated as proportional to the total pressure The proportion found by Paneth (1937), 0·03 % by volume was used.

Data on absorptions

The absorption of solar energy in oxygen has only been taken into account for wave-lengths longer than 1750 A It appears to be of little importance below 55 km. but shows signs of becoming important at greater heights If calculations at such heights are attempted the effects of the strong absorption band at 1450 A. should be included. Gotz (1936) has made some calculations of the absorption in oxygen.

He also gives a table showing total absorbed solar energy of wave-length less than 3400 A. in 0·200 cm. of ozone The values are in fair agreement with the solar totals obtained during the calculations, though the conditions are not identical The absorption coefficients for oxygen used in this paper are given by Granath (1929)

The absorption coefficients for ozone are in the ultra-violet from Fabry & Buisson (1913), Lambert, Dejardin & Chalonge (1927), Lauchli (1929); in the visible from Colange (1927), and in the infra-red from Summerfield (1941). Summerfield has verified that the pressure effect is important only in the infra-red, and the necessary corrections have been made for the main band at 9 6μ from his thesis, on the basis of a band width of 1 0μ

Elsasser (1942) concludes that the half-width of the individual lines comprising an infra-red band is proportional to the square root of the total pressure, and that for water vapour the lines are still without much overlap at atmospheric pressure A blend of the coefficients of Fowle (1917) and Hettner (1918) has been used in all previous calculations As a check the amounts of water vapour in the layers were considered as at atmospheric pressure, and the emissivity calculated at 20° C The emissivities as a percentage of black were compared with the observed results of figure 29 in Elsasser's monograph, wherever the amounts of water vapour coincided The agreement was good, and the absorptions which gave this agreement were used in one set of temperature calculations, which are therefore not adjusted for the pressure effect Subsequently another set of temperature calculations was completed using the absorptions adjusted for the pressure at the lower boundary of the layer concerned

For the upper layers an extrapolation was necessary, and for this purpose a replot on log-log paper was made. The curve is slightly concave downwards, but for the extrapolation a straight line was used to give the absorptions for the upper layers. These are therefore somewhat too high, but the pressure correction makes their influence on the calculated temperature rather small In cases where water vapour is not corrected for pressure the 10 and 40 % values of water vapour would be slightly low for the upper three layers.

The emission of the 14μ band (whose width is 2·0μ) for carbon dioxide was taken from the curve on p. 80 of Elsasser's monograph The curve was replotted on log-log paper which makes it concave downwards For the small quantities involved in the upper layers an extrapolation is necessary. This was done on the basis of a straight line, which gives too large absorptions and emissions Even so the carbon dioxide contribution, whether corrected or not corrected for the pressure effect by the square-root law, turned out to be negligible above 30 km. There may be a certain overlapping in absorption and emission for carbon dioxide and water vapour about 15 to 16μ This has been neglected, each constituent being treated individually.

There is an effect of temperature on absorption. It is smaller than the pressure effect, the numerical values are uncertain, and it adds an extra complication to the successive approximation scheme used In all cases calculated this temperature effect has been neglected.

Data on energy entering the ozonosphere

The solar energy outside the atmosphere in the region of ozone absorption is not known with great accuracy. Coblentz & Stair (1936, 1943), Stair & Hand (1939), Pettit (1940) and Hunter (1943) have studied and discussed the matter Some results give 4000° K, most 6000° K, and some even higher temperatures for the region around the short-wave limit of the solar spectrum Most of the sets of calculations were done on the basis of 6000° K, but for comparison two are included, based on the lower figure.

The combined radiation of the earth's surface and the troposphere is required for E_λ at 11 km This has been computed for the many wave-lengths necessary, using observed average values of temperature, humidity and ozone concentration (Napier Shaw 1928).

Results

In table 2 are summarized the conditions governing each separate calculation, and the temperatures obtained for each of the nine layers, with an estimated error of about 3° K. The temperatures given in plain figures were obtained with ozone and carbon dioxide absorption corrected for pressure in the infra-red, but water vapour not so corrected. The temperatures in italics were obtained with the water vapour corrected for the pressure effect This pressure correction is large, particularly in the upper layers, but it affects both sides of the equation, and so the temperature differences are not greater than 20 %.

Table 2. Calculated air temperatures in °K

season	summer	summer	summer	summer	winter	summer
amt O_3	0 280	0 280	0 280	0 200	0·280	0 280
amt H_2O (%)	0	10	40	10	10	10
solar temp (°K)	6000	6000	6000	6000	6000	4000
ht in km						
50–55	452	415 *448*	344 *441*	410 *445*	406 *439*	323 *347*
45–50	429	410 *424*	361 *421*	409 *422*	364 *375*	321 *333*
40–45	399	385 *397*	350 *394*	382 *390*	305 *314*	311 *319*
35–40	335	324 *332*	295 *327*	320 *330*	278 *285*	291 *301*
30–35	296	285 *295*	262 *291*	281 *292*	262 *273*	272 *282*
25–30	275	258 *272*	244 *265*	257 *269*	246 *256*	252 *266*
20–25	254	241 *249*	240 *239*	239 *247*	230 *238*	236 *245*
15–20	239	232 *232*	232 *221*	229 *225*	221 *221*	229 *229*
11–15	228	218 *217*	211 *209*	215 *208*	208 *209*	217 *215*

Most of these data are plotted in figure 1 where the pressure correction has been applied, and in figure 2 where the water-vapour absorption is uncorrected for the pressure effect. Figure 2 therefore can be compared with previous results to show the influence of the new data on the infra-red absorption of ozone. For convenience

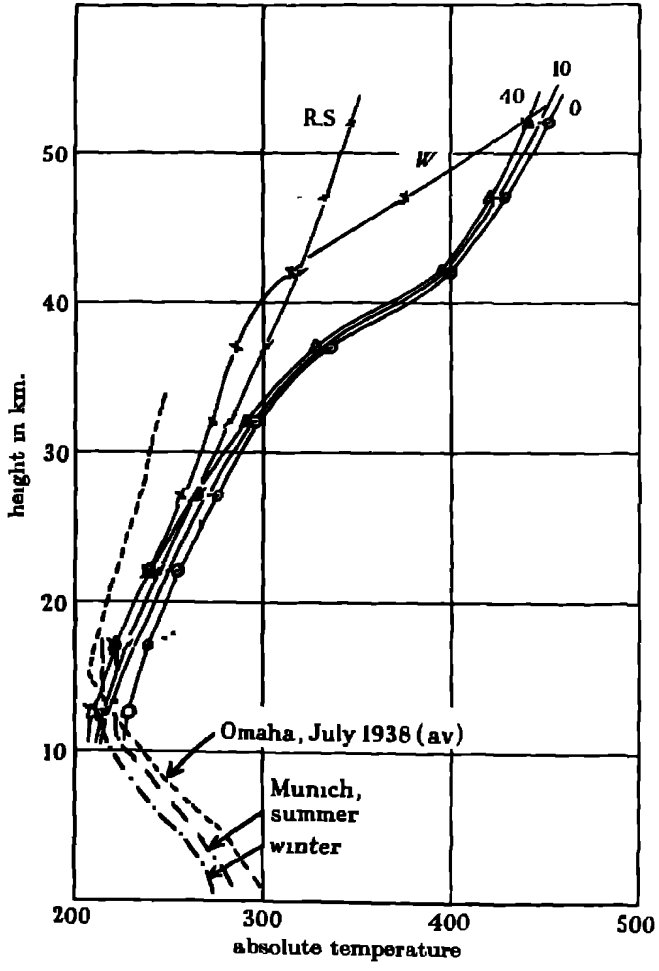


FIGURE 1 All absorptions corrected for pressure

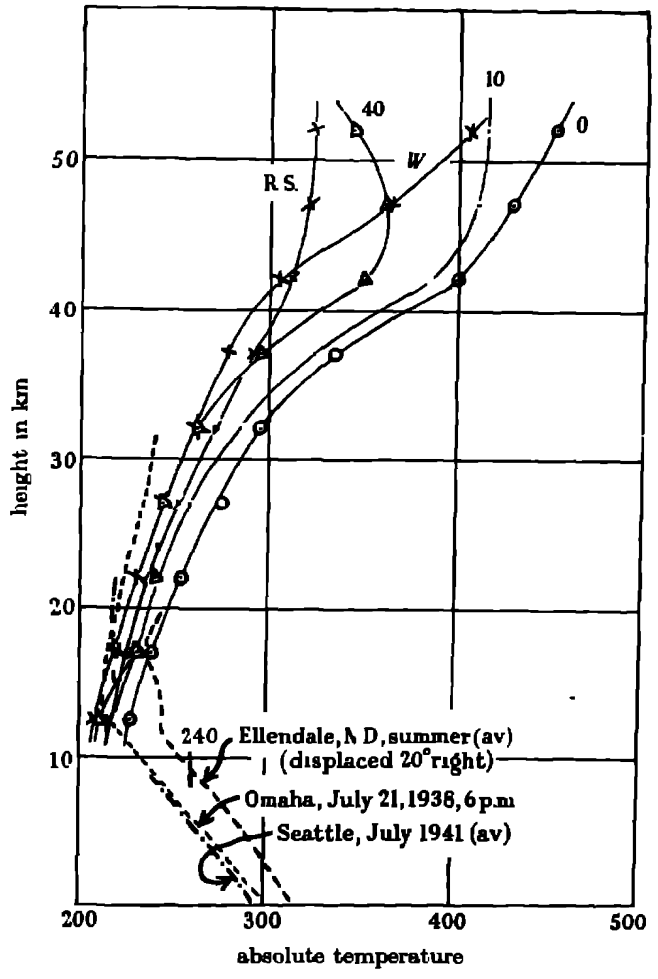


FIGURE 2. H_2O absorption not corrected for pressure.

in comparing the two figures it should be noted that the curve marked zero in each is the same, and represents the limit in which water vapour is negligible above 11 km.

Discussion

In each figure are given some temperature results obtained with sounding balloons. The values for Munich (latitude 48° N) are an average for summer, no time of the flights being specified. The Seattle (latitude 47° N) values are taken from the *U S Monthly Weather Review*, and refer to flights starting about 9 30 p m. local standard time. Two curves are shown for Omaha (latitude 41° N), one an average for July 1938, and the other a single flight on 21 July 1938 All flights started late in the day, about 6 p m The curve for Ellendale (latitude 46° N) is displaced 20° to the right to avoid overlapping (Lennahan 1938) None of these curves is strictly comparable with the calculated values, since the attainment of radiation equilibrium at midday has been assumed for the calculations. Failure to attain this equilibrium, or cooling from it by 6 p m or by 10 p m , would certainly make the calculated temperatures higher than those observed.

Ballard (1941) obtains a cooling of 10° K from 6 p.m to 2 a m at the height of 25 km., based on a small number of observations. If a cooling of 15° K could be assumed from midday to midnight at 25 km., the observed temperatures would nearly agree with the calculated values This agreement may imply that even below 30 km the diurnal temperature changes are rapid enough to allow the attainment of radiation equilibrium.

The difference between observations over Seattle on the west coast and over Omaha in the great plains is very similar to the difference in temperature distributions for temperate and tropical latitudes, both in height of tropopause and lapse rate above it (Chiplonkar 1940, Jaumotte 1937) Hafer (1940) states that at Omaha in general the summer shows a large temperature inversion and the winter months little or none in the stratosphere. The same effects might be observed at greater heights over Ellendale

It will be noticed that the pressure corrected curves for zero, 10 and 40 % water vapour are very close together above 25 km , while this is not the case for the uncorrected ones. The crowding towards the zero curve in the higher layers is due to the fact that the pressure correction is equivalent to decreasing greatly the effective amount of water vapour in these layers.

A significant difference is evident from summer to winter below 25 km. The curves, however, show about 10° K difference near 20 km., and this is in agreement with balloon observations (Munich and Seattle) The calculations for winter conditions do not take into account any possible change from summer in the water-vapour content of the ozonosphere or in the vapour pressures in the troposphere Another feature of the winter curves is the small change from summer in the temperature of the 50 to 55 km layer. This is mainly due to the large zenith angle of the sun and the consequent long path of the parallel solar rays. For the shorter wave-lengths of the

ultra-violet ozone band the absorption is complete in this one layer, almost compensating for the greater area over which the energy is spread. In the next few layers down to 35 km. this depletion of energy in the top layer does not leave enough to compensate for the larger area, and the temperature differences are large. Still lower, below 30 km , the other terms in the equation become more important than the solar term and the temperature differences become small again.

The curves for reduced solar energy show that the temperature could still be above 300° K around 50 km. in spite of the 2000° K drop in the effective temperature for solar ultra-violet.

The curves for 0·200 cm of ozone have not been plotted. They show no change above 40 km as would be expected, since the ozone decrease is in the layers below this height due largely to inaccuracy of determination above 40 km. A real decrease of ozone and temperature almost certainly occurs here In the middle layers there is a tendency for the solar energy to be absorbed lower down (in contrast to the winter calculations) and a small decrease in temperature is shown from 15 to 30 km. This change is less than of that expected by Dobson (1942)

Vassy & Vassy (1939) infer that the mean temperature of atmospheric ozone is low, but nevertheless there is a temperature increase in the upper part of the ozonosphere. The calculated results of the present paper agree, since the bulk of the ozone is below 30 km where the average calculated temperature is about 240° K. Regener (1941) discusses the high temperature of the upper ozonosphere, and its explanation by the absorption of solar energy. He suggests that spectral methods may yet be used to determine the temperature distribution at great heights.

Conclusion

The results generally indicate that absorption of solar radiation by ozone will explain the existence of high temperatures around 35 to 50 km , though the ozone maximum is at 22 km The small increase in the lower stratosphere also agrees in slope with the Omaha balloon observations (which may not be representative of the earth as a whole (Wigand 1931)), but the calculated temperatures are 15 to 20° K too high at 25 km. The assumptions regarding water vapour are reasonable, but in truth there may be irregular variations with height about which nothing is now known. The approximate agreement with directly observed temperatures serves to justify the primary assumption of radiation equilibrium as a good working hypothesis For heights below 30 km this aspect of the problem should repay further study, including an extensive analysis of balloon observations under widely different geographical and climatic conditions.

References

Ballard 1941 *Mon. Weath. Rev. Wash.* **69**, 33.
Brunt & Kapur 1938 *Quart. J. R Met. Soc.* **64**, 510.
Chiplonkar 1940 *Proc. Ind. Acad. Sci.* A, **11**, 39.
Coblentz & Stair 1936 *Bur. Stand. J. Res , Wash.*, **17**, 1
Coblentz & Stair 1943 *Bur. Stand. J. Res., Wash* , **30**, 434
Colange 1927 *J. Phys. Radium*, p 254.
Cowling 1943 *Rep. Progr. Phys.* **9**, 29
Dobson 1942 *Quart. J. R. Met Soc.* **68**, 202.
Elsasser 1942 *Harv. Met. Stud.* no 6, p. 96.
Fabry & Buisson 1913 *J Phys. Radium*, **5**, 196
Fowle 1917 *Smithson Misc Coll.* **68**, No 8.
Gotz 1936 *Quart. J. R. Met. Soc.* **62** (Suppl), 41.
Gotz, Meetham & Dobson 1934 *Proc. Roy. Soc.* A, **145**, 416.
Gowan 1928 *Proc. Roy Soc* A, **120**, 655.
Gowan 1930 *Proc. Roy. Soc* A, **128**, 531
Gowan 1936 *Quart. J R. Met. Soc.* **62** (Suppl.), 34.
Granath 1929 *Phys. Rev* **34**, 11, 1045.
Hafer 1940 *Mon. Weath. Rev. Wash* **68**, 125.
Hettner 1918 *Ann Phys., Lpz.*, **55**, 476.
Hunter 1943 *Rep. Progr Phys* **9**, 5.
Jaumotte 1936 *Bruxelles, Mem. R. Inst Met.* 5.
Jaumotte 1937 *Beitr. Geophys* **50**, 403.
Lambert, Dejardin & Chalonge 1927 *Bull Obs. Lyon*, **9**, 45
Lauchli 1929 *Z. Phys* **53**, 92
Lennahan 1938 *Mon. Weath. Rev. Wash* Suppl. **38**.
Paneth 1937 *Quart. J. R Met. Soc.* **63**, 433.
Pettit 1940 *Contr. Mt Wilson Obs.* **624**.
Pettit 1940 *Astrophys J* **91**, 159.
Price 1943 *Rep. Progr Phys* **9**, 10.
Regener 1941 *Naturwissenschaften*, **29**, 479.
Shaw, N. 1928 *Handbook of Meteorology*
Stair & Hand 1939 *Mon. Weath Rev. Wash* **67**, 331.
Strong 1939 *Phys Rev.* **55**, 1114
Summerfield 1941 Doc. Thesis. Cal Tech
Sutherland & Callendar 1943 *Rep. Progr. Phys* **9**, 18.
Vassy & Vassy 1939 *C.R Acad. Sci , Paris*, **208**, 1518, 1664.
Wigand 1931 *Beitr. Phys Frei Atmos* **17**, 288.
Wigand 1939 *Quart J. R. Met. Soc.* **65**, 321.

Night cooling of the ozonosphere

By E. H. Gowan, *University of Alberta, Edmonton*

(*Communicated by G. M. B. Dobson, F.R.S.—Received* 24 *September* 1946 —*Read* 21 *November* 1946)

Some estimate of the cooling of the ozonosphere can be made using the calculated solar energy absorbed in various layers After sunset this source of energy is removed and cooling occurs mainly by radiation up and down The conclusion is reached that during a night the cooling is about 30° C at 50 km. and only a fraction of a degree below 25 km This means that the ozonosphere temperature inversion around 40 km will persist throughout the night

Ballard (1941) finds that at about 25 km height during July over Nebraska the air temperature drops by about 10° C during the eight hours from 6 p.m. to 2 a.m , and ordinarily by less than 2° C around 15 km. On the other hand, Penndorf (1936) obtained a much smaller diurnal variation of temperature in the stratosphere, the maximum being 0·3° C per day around 50 km. height.

The amount of radiation energy being absorbed by various layers of the ozonosphere was obtained by Gowan (1936, 1947) during calculations of temperature on the assumption of radiation equilibrium At sunset one source of this energy is cut off but the others remain. The temperature should therefore fall during the night, and an estimate of how much can be made with the numerical values available. The results which follow lie generally between the two sets of data referred to above, for comparable heights

The assumptions of two conditions—radiation equilibrium before sunset, and that sunset is instantaneous—gives a preliminary basis for calculation Let the mass of air in a 1 sq cm. column of any layer be m g ; let the specific heat at constant pressure be c_p, let R be the net radiation from the layer in g cal /sq cm./hr., and let dT degrees be the fall in temperature in dt hours. Then the conditions are represented by the equation

$$-mc_p dT = 2R dt.$$

There is no net exchange of radiation horizontally in any layer, so $2R$ is the total net radiation from a sample column This may be considered constant as a first approximation, or may vary according to several plausible laws. Below are outlined five different ways of treating this net radiation.

(1) $2R$ = a constant, namely, the solar energy which was being absorbed, and after sunset must be radiated, no account being taken of the change in temperature of the layer. Thus $\partial T/\partial t = -2R/mc_p$; and $t(\partial T/\partial t)$ gives the temperature change in t hours.

(2) $2R = \beta T$, where β is a constant, and T is the absolute temperature of the layer t hours after sunset. Thus $-mc_p dT = \beta T dt$, which gives on integration

$$\log_{10} T = \log_{10} T_0 - \frac{0{\cdot}434\beta t}{mc_p}.$$

(3) $2R = \beta(T - T_s)$, where T_s is the estimated effective temperature of the other adjacent layers. This is Newton's law of cooling if T_s is constant, as here provisionally assumed. The actual decrease in T_s during the night is small because it mostly represents the lower layers containing the bulk of the absorbing material, and these are found to change relatively little in temperature. The factor $(T - T_s)$ tends to remain large if T_s decreases with T, and the calculated cooling would be larger than if T_s is constant. By considering T_s constant some compensation is obtained for the maximizing effect of the assumptions of instantaneous sunset with midday equilibrium temperature. By substitution $-mc_p dT = \beta(T - T_s)\,dt$, which gives

$$\log_{10}(T - T_s) = \log_{10}(T_0 - T_s) - \frac{0{\cdot}434\beta t}{mc_p}$$

on integration

$$\beta = \frac{2R}{(T_0 - T_s)}.$$

(4) $2R = \beta(T - T_\infty)^n$, where T_∞ is the equilibrium temperature for night, i e. with no solar energy reaching the layer Substituting in the original equation, and using $\beta = 2R/(T_0 - T_\infty)^n$ gives

$$(T - T_\infty)^{n-1} = \frac{mc_p(T_0 - T_\infty)^n}{(n-1)\,2Rt + mc_p(T_0 - T_\infty)}$$

on integration and simplification This formula was applied in two cases: (a) $n = 2$, (b) $n = 4$.

The data used were from the calculation of temperature under radiation equilibrium for midsummer at latitude 50 with 0 280 cm of ozone, and 10 % relative humidity at the tropopause The absorption of all constituents was corrected for pressure The cooling in 8 hr was found by each of the five methods outlined above. The results are summarized for comparison in table 1 The largest values come from method (1) and the smallest from (4b). The variations, however, are less significant than the general agreement of the order of magnitude, tens of degrees above 35 km. and fractions of a degree below 25 km

TABLE 1 DROP IN TEMPERATURE (°C)

	8 hr					16 hr
ht in km.	(1)	(2)	(3)	(4a)	(4b)	(4b)
50-55	23	22	21	21	20	30
45–50	32	30	28	26	24	26
40–45	32	31	27	27	23	13
35–40	13	13	12	12	10	6
30–35	3·1	3 1	3 0	3 0	3·0	2 0
25–30	1 0	1 0	0 9	1·0	1 0	0 9
20–25	0 3	0 3	0·3	0 3	0 3	0·3
15–20	0·1	0 1	0 1	0·1	0·1	0 1
11–15	0 1	0·1	0·05	0·03	0 03	0 03

Discussion

The actual course of the temperature change will be a small, slow cooling from shortly after midday to sunset, followed by a more rapid cooling from sunset to sunrise. The night rate will be smaller than the smallest in the table, but with the start of cooling during the afternoon the final result from, say, 1 p m. to sunrise may not be very different from the values in the table.

From sunrise to about 1 p.m. the ozone layers are warming up again. During spring the period of warming is lengthening, and that of cooling shortening, producing an all-round increase in temperature. During autumn the reverse process takes temperatures down to their winter values The cooling during the 16 hr of a midwinter night was calculated by (4*b*) This was plotted along with the corresponding summer cooling for comparison as shown in figure 1. Maximum absorption of solar energy in ozone occurs lower down in summer when the sun's rays have a short path through the atmosphere. This gives more energy to be radiated after sunset, and hence a maximum in the estimated cooling around 45 km The winter curve would probably show a similar maximum at some greater height than 55 km where the ozone concentration is small and unknown, making calculation impossible Möller (1935) has calculated the heat balance of the free atmosphere over Lindenberg in June up to 10 km The result at 10 km is about 0·5° C cooling per day, or about 0 2° C per 8 hr., and if extrapolated to 13 km. would be less than 0·1° C per day. Tanck (1940) calculated the heating effect due to absorption of sunlight, and found a similar value. Both these observers agree that the temperature change decreases with height in the high troposphere, and then increases again in the upper stratosphere. Elsasser (1942) also discusses this question and comes to a similar conclusion through the use of his radiation chart.

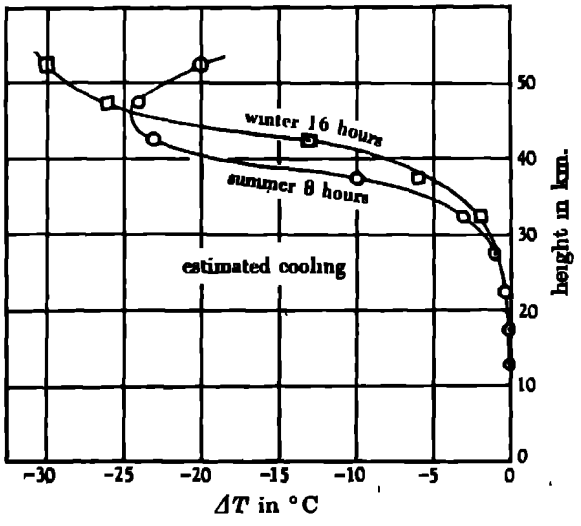


Figure 1. Graph of ozonosphere cooling.

Baur & Phillips (1934, 1935, 1936) have studied the heat and radiation balance in the atmosphere. This was done for clear skies, and also including the effects of turbidity and cloudiness, for summer, winter and equinoctial seasons. The distribution with height is quite different with and without cloud, but the values agree at about 10 km., and are substantially the same as Möller's figure.

Barbier, Chalonge & Vassy (1939) discuss the spectroscopic evidence for a low temperature of the ozonosphere at night. The method is based on the contrast in the absorption bands of ozone, and naturally refers to the region where the bulk of the ozone is found, i.e. below 30 km Vassy & Vassy (1939) also discuss the question, and show that the day and night readings are not different within the precision of measurement At greater heights Bartels (1936) has estimated that the change in temperature between day and night around 200 km is certainly less than 100° C

All this auxiliary evidence is confirmatory, or at least not contradictory, within its range. The calculations in this paper show less than 2° C per night under 30 km , yet only up to 30° C per night around 50 km The direct evidence of a temperature change from day to night is sparse and contradictory. It seems from the work of Ballard (1941) that the insolation error during daytime flights is not large, but it may be important when the differences are required for evaluation of the cooling. There are known to be large and relatively sudden changes in the height of the tropopause, and consequently in the temperature of the lower stratosphere These have no connexion with day and night, but are the result of air mass movements which seem to extend some way into the stratosphere The great discrepancies which appear in the average observations of different observers, are mainly due to geography and climate. It seems certain that the cooling of the lower stratosphere, perhaps up to 30 km , is not governed by radiation. Above 30 km. it is more probable that radiation plays a dominant role in any diurnal temperature changes, though winds have been observed during observations of noctilucent clouds and meteor trails.

If the cooling of the lower stratosphere depends on convection and advection it follows that radiation equilibrium may never be attained, as assumed in the original calculations (Gowan 1947). For those it was assumed that there is sufficient mixing to keep the composition nearly constant, but not enough convection to interfere with the attainment of radiation equilibrium. The agreement between calculated temperatures and balloon observations is good for the lowest layer, 11 to 15 km., but around 15 to 25 km. is only fair, the calculated temperatures being 10° C too high. This disagreement might be due to any one or all of the following factors: (1) Radiation equilibrium may not be attained, hence the observed temperature would never be as high as the calculated temperature. (2) Insufficient water vapour may have been assumed for these layers, since for higher amounts the calculated temperatures are lower. (3) If more carbon dioxide were assumed its extra radiation would lead to a lower calculated temperature. (4) The relatively weak ozone band at 14μ was neglected. This overlaps with both carbon dioxide and water vapour but could have more effect from 15 to 25 km. than from 11 to 15 km., due to the

maximum concentration of ozone around 22 km. Some adjustment among these factors might be found by trial which would give agreement at all heights observed directly It seems likely that more observations of composition at greater heights may soon be available from experiments with rockets. Then perhaps fewer trial calculations will be necessary to reach an agreement over all the layers within the ozonosphere.

Although these results are tentative it seems clear that even the temperature change above 35 km. is not sufficiently great to destroy the ozonosphere temperature inversion during a summer or a winter night. Hence the existence of a temperature inversion is a valid explanation of the minimum in number of meteor disappearances, and of the abnormal audibility of explosions extending throughout the night

References

Ballard 1941 *Mon. Weath. Rev Wash* **69**, 33.
Barbier, Chalonge & Vassy 1939 *J. Phys. Radium*, **10**, 550.
Bartels 1936 *Z. Geophys* **12**, 368.
Baur & Phillips 1934 *Beitr. Geophys* **42**, 160
Baur & Phillips 1935 *Beitr. Geophys.* **45**, 82
Baur & Phillips 1936 *Beitr Geophys.* **47**, 218
Elsasser 1942 *Harv Met. Stud* no 6, p. 96.
Gowan 1936 *Quart J.R. Met Soc.* **62** (Suppl.), 34.
Gowan 1947 *Proc. Roy. Soc* A, **190**, 219
Möller 1935 *Met. Z.* **52**, 408.
Penndorf 1936 *Veröff. Geophys. Inst. Univ., Lpz.*, **8**, 181.
Tanck 1940 *Ann Hydrogr , Berl ,* **6**, 47
Vassy & Vassy 1939 *C R. Acad. Sci., Paris*, **208**, 1518, 1664.

Turbulence and diffusion in the lower atmosphere with particular reference to the lateral effect

By D. R. Davies, Ph.D.

(*Communicated by L. Rosenhead, F.R.S.—Received* 29 *July* 1946)

It has been established experimentally by Pasquill that the theory of atmospheric turbulence put forward by O. G. Sutton in 1934 leads to satisfactory theoretical rates of *evaporation* from plane, saturated, liquid surfaces. The theory, however, overestimates the vapour concentration at any point and underestimates the rate of decay of concentration with height above such evaporating surfaces of *finite* lateral dimensions In the present paper it is suggested that one of the reasons for this discrepancy lies in the neglect of the lateral component of turbulence The aim of the present investigation is to estimate quantitatively the effect on the values of the theoretical vapour *concentration* produced by the introduction of a lateral component of turbulence

It is shown that the theoretical evaporation is not appreciably affected by the modification introduced but that the concentration of vapour, and the rate of decay of concentration with height, are much nearer the experimental values, which, however, are very scanty The investigation deals with a saturated area of parabolic shape Extensions of the theory are contemplated

1. Introduction

It has been established, as a result of a long series of experiments by Pasquill (1943), that the theory of atmospheric turbulence put forward by O. G. Sutton (1932, 1934) leads to satisfactory theoretical rates of *evaporation* from plane, saturated, liquid surfaces. The theory, however, overestimates the vapour *concentration* at any point and underestimates the rate of decay of concentration with height above any such evaporating surfaces of *finite* lateral dimensions In the present paper it is suggested that one of the reasons for this discrepancy lies in the neglect of the lateral component of turbulence. The special case of concentration of vapour above a saturated liquid surface of *parabolic* shape is considered.

In addition to work by O. G. Sutton and Pasquill, referred to above, problems in this field have been attacked by W. G L. Sutton (1943), and Frost (1946), but in all these expositions only the dispersion of liquid vapour by eddy diffusion *normal* to the evaporating surface was taken into account

A theory of turbulence has yet to be formulated which will give rigorous expressions for the dependence of eddy diffusivity on height above ground. The result is that empirical formulae have had to be adopted by these authors to give the variation of wind speed and normal diffusivity with height These expressions on substitution in the diffusion equation (2·2) yield the differential equation

$$\frac{\partial \chi}{\partial x} = \text{const } z^{-m} \frac{\partial}{\partial z}\left(z^{1-m} \frac{\partial \chi}{\partial z}\right), \tag{1·1}$$

where χ is the concentration at a height z and at a distance x along the wet strip. This equation has been solved, by O G. Sutton, W G L Sutton and Frost, subject to certain prescribed boundary conditions (specified later in § 2). The calculated rates of *evaporation* derived on this hypothesis gave good agreement with experi

mental results for evaporation from plane areas of various shapes (Pasquill 1943) The agreement is not unexpected, because near the liquid surface it is *clearly* the *diffusivity* normal to the surface which is the *operative agent* in transporting the vapour from the surface, except for a very small effect near the bounding edges

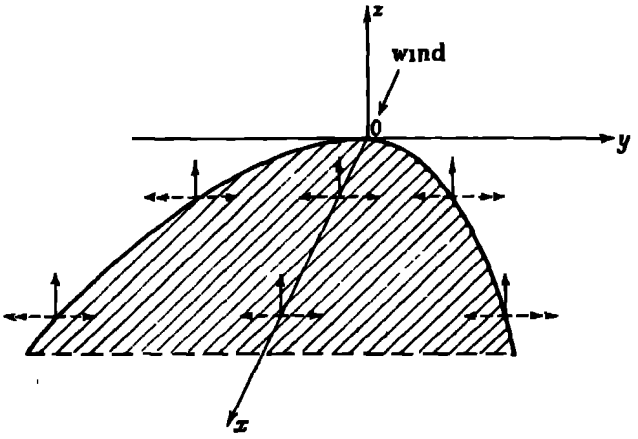


FIGURE 1

Let x, y, z be orthogonal space co-ordinates with origin at the vertex of a parabolic saturated area lying in a horizontal plane, as in figure 1; z is measured vertically upwards and x along the axis of the area. In the theory established by O G. Sutton (1934) transfer of mass across the planes z = constant is considered The present investigation introduces additional transfer across the planes y = constant. Evaporation, however, is clearly given by the transfer normal to the plane $z = 0$, indicated by full arrows in figure 1; transfer of mass across the planes y = constant at $z = 0$, indicated by the single dotted arrows in the diagram, will not have any influence on evaporation. There may, however, be an appreciable effect near the bounding edges, indicated by double dotted arrows. Hence the easily computable expressions for *evaporation*, obtained by neglecting lateral diffusion, are adequate.

The distribution of vapour in the air stream at finite distances *above* and down-wind of the surface is dependent on both the vertical and lateral diffusivities, and, as might have been expected on physical grounds, the theory dependent on one-dimensional dispersion underestimates the decay of vapour concentration at points above and down-wind of areas of finite *lateral* dimensions, except where the latter is *very large* (Frost 1946)

The first successful attempt at a solution involving eddy diffusion in two dimensions was given by Dr W. G Bickley (unpublished) for an infinite quadrant, with one bounding edge normal to the mean wind direction He assumed that wind speed and diffusivity was constant with height above the surface. It was pointed out to the present author by Dr Bickley that a solution could also be obtained for a parabolic area (with vertex pointing up-wind and axis orientated in the mean wind

direction), again on the assumption that wind speed and diffusion were constant. The method was worked out in detail by the author and is included in § 3 of this paper.

In § 4 an empirical expression, for the dependence of lateral diffusivity on height, is proposed which, together with the expression for normal diffusivity and wind speed used by O G. Sutton (1934), defines a practical model of the lower atmosphere in which wind speed, normal and lateral diffusivity vary as simple powers of height above the surface. Substitution of these coefficients in equation (2·2) yields the new differential equation (5·2) which, together with prescribed boundary conditions, presents mathematical difficulties which have been resolved as yet only for a saturated area of parabolic shape. This solution is given in § 5. Its value lies in its comparison with theoretical results derived from solutions of equation (1·1) and with experimental results. in this way the effects of neglecting the lateral dispersion may be estimated

2. The equation of diffusion and boundary conditions

If U is the mean wind velocity at height z, in the positive sense of the x co-ordinate, and K_x, K_y, K_z are the coefficients of diffusivity in x, y and z directions respectively, the equation satisfied by the concentration $\chi(x, y, z)$ of vapour is

$$\frac{\partial \chi}{\partial t} + U\frac{\partial \chi}{\partial x} = \frac{\partial}{\partial x}\left[K_x\frac{\partial \chi}{\partial x}\right] + \frac{\partial}{\partial y}\left[K_y\frac{\partial \chi}{\partial y}\right] + \frac{\partial}{\partial z}\left[K_z\frac{\partial \chi}{\partial z}\right]. \tag{2·1}$$

Here K_x, K_y and K_z are such that the rates per unit area at which χ is transferred across the planes $x, y, z =$ const. in the positive directions are

$$-K_x\frac{\partial \chi}{\partial x}, \quad -K_y\frac{\partial \chi}{\partial y} \quad \text{and} \quad -K_z\frac{\partial \chi}{\partial z}.$$

Steady state conditions are assumed and K_x is neglected, so that the diffusion equation (2·1) becomes

$$U\frac{\partial \chi}{\partial x} = \frac{\partial}{\partial y}\left[K_y\frac{\partial \chi}{\partial y}\right] + \frac{\partial}{\partial z}\left[K_z\frac{\partial \chi}{\partial z}\right]. \tag{2·2}$$

It is assumed that the vapour pressure of the evaporating liquid approaches its saturation value at the temperature of the surface as the surface is approached. This implies that the vapour concentration χ assumes a constant value χ_0 at the surface of the saturated area.

The appropriate boundary conditions then are

(i) $\dfrac{\chi}{\chi_0} \to 1$, as $z \to 0$ for values of x and y *inside* the saturated area,

(ii) $\lim\limits_{z\to 0}\left\{K_z\dfrac{\partial \chi}{\partial z}\right\} \to 0$ for x and y *outside* the saturated area,

(iii) $\chi \to 0, \quad x > 0, \quad z \to \infty,$

(iv) $\chi \to 0, \quad y \to \pm\infty, \quad z \geqslant 0,$

(v) $\chi \to 0, \quad x \to +0, \quad z > 0.$

3. Wind speed and diffusivities constant with height above evaporating surface

It is convenient to substitute non-dimensional variables given by

$$x = \xi\alpha U, \quad y = \eta\sqrt{(\alpha K_y)} \quad \text{and} \quad z = \zeta\sqrt{(\alpha K_z)}, \tag{3·1}$$

where α is a scale factor of the dimensions of time and U, K_y, K_z are constants. In these co-ordinates equation (2 2) becomes

$$\frac{\partial\chi}{\partial\xi} = \frac{\partial^2\chi}{\partial\eta^2} + \frac{\partial^2\chi}{\partial\zeta^2}. \tag{3·2}$$

The initial area $y^2 = 4Ax$ becomes, in the new co-ordinates, $\eta^2 = 4\mu\xi$, where $\mu = U/K_yA$.

Now suppose that space is mapped out by a system of parabolas given by

$$X = \xi, \quad Y = \frac{\eta}{2\sqrt{(\mu\xi)}} \quad \text{and} \quad Z = \frac{\zeta}{2\sqrt{(\mu\xi)}}. \tag{3·3}$$

Equation (3·3) becomes

$$X\frac{\partial\chi}{\partial X} = \tfrac{1}{2}Y\frac{\partial\chi}{\partial Y} + \tfrac{1}{2}Z\frac{\partial\chi}{\partial Z} + \frac{1}{4\mu}\left[\frac{\partial^2}{\partial Y^2} + \frac{\partial^2}{\partial Z^2}\right]\chi. \tag{3 4}$$

A solution for χ independent of X clearly exists, and this is consistent with the boundary conditions The equation for χ depending on Y and Z then is

$$Y\frac{\partial\chi}{\partial Y} + Z\frac{\partial\chi}{\partial Z} + \frac{1}{2\mu}\left[\frac{\partial^2}{\partial Y^2} + \frac{\partial^2}{\partial Z^2}\right]\chi = 0, \tag{3 5}$$

with the appropriate boundary conditions

(i) $\dfrac{\chi}{\chi_0} \to 1, \quad Z \to 0, \quad -1 \leqslant Y \leqslant 1,$

(ii) $\dfrac{\partial\chi}{\partial Z} = 0, \quad Z \to 0, \quad 1 \leqslant Y \leqslant -1,$

(iii) $\chi \to 0, \quad Z \to +\infty,$

(iv) $\chi \to 0, \quad Y \to \pm\infty.$

These equations suggest the further transformation into elliptic co-ordinates

$$Y + iZ = c\cosh(\beta + i\theta). \tag{3·6}$$

Since the area is bounded by

$$\eta^2 = 4\mu\xi, \text{ then } c = 1.$$

Equation (3·5) becomes

$$\frac{1}{2\mu}\frac{\partial^2\chi}{\partial\beta^2} + \sinh\beta\cosh\beta\frac{\partial\chi}{\partial\beta} + \frac{1}{2\mu}\frac{\partial^2\chi}{\partial\theta^2} - \sin\theta\cos\theta\frac{\partial\chi}{\partial\theta} = 0, \tag{3·7}$$

with the boundary conditions

(i) $\frac{\chi}{\chi_0} \to 1$ as $\beta \to 0$, $0 < \theta < \pi$,

(ii) $\frac{\partial \chi}{\partial \theta} = 0$ for $\theta = 0$ and π, $\beta > 0$,

(iii) $\chi \to 0$ as $\beta \to \infty$, $0 \leqslant \theta \leqslant \pi$.

A solution independent of β exists of the form

$$\chi = K_1 + K_2 \int_0^{\beta} \exp[-\mu \sinh^2 \phi]\, d\phi$$

This may be adjusted to satisfy the conditions, if $K_1 = 1$ and $K_2 = -1/I(\mu)$,

where
$$I(\mu) = \int_0^{\infty} \exp[-\mu \sinh^2 \phi]\, d\phi$$
$$= \tfrac{1}{2} \exp(\tfrac{1}{2}\mu)\, K_0(\tfrac{1}{2}\mu),$$

and K_0 is the modified Bessel function of the second kind with imaginary argument. So that the equation for χ is

$$\frac{\chi}{\chi_0} = 1 - \frac{1}{I(\mu)} \int_0^{\beta} \exp[-\mu \sinh^2 \phi]\, d\phi. \tag{3·8}$$

The rate of evaporation at any point (x, y) is given by

$$E(x, y) = -\lim_{z \to 0} \left\{ K_z \frac{\partial \chi}{\partial z} \right\}, \tag{3·9}$$

and from the transformations it follows easily that

$$E(x, y) = \sqrt{\left(\frac{K_z U}{\mu x}\right)} \frac{\chi_0}{\exp(\tfrac{1}{2}\mu)\, K_0(\tfrac{1}{2}\mu)} \frac{1}{(\sin\theta)_{\beta=0}}, \tag{3·10}$$

by using the relation
$$\left(\frac{\partial \chi}{\partial Z}\right)_{Z=0} = \frac{1}{\sin\theta} \left(\frac{\partial \chi}{\partial \beta}\right)_{\beta=0}$$

At $\beta = 0$, i.e on the saturated strip,

$$\cos\theta = Y = \frac{\eta}{2\sqrt{(\mu\xi)}} = \frac{y}{2\sqrt{(Ax)}},$$

giving
$$\sin\theta = \frac{\sqrt{(4Ax - y^2)}}{2\sqrt{(Ax)}}.$$

Substitution for $\sin\theta$ in (3·10) gives

$$E(x, y) = 2\sqrt{(K_y K_z)} \frac{\chi_0}{\exp(\tfrac{1}{2}\mu)\, K_0(\tfrac{1}{2}\mu)\, (4Ax - y^2)^{\frac{1}{2}}}. \tag{3·11}$$

The total rate of evaporation from an area of down-wind length x is given by integration of equation (3 10) with respect to θ from 0 to π and from 0 to x, down-wind, to give

$$E = \chi_0 \frac{2\pi\sqrt{(K_y K_z)}\,x}{\exp(\frac{1}{2}\mu)\,K_0(\frac{1}{2}\mu)}. \tag{3·12}$$

4. Specification of the diffusion coefficients as functions of the degree of turbulence and of height above the surface

The formulae adopted to give the functional dependence of mean wind speed and of vertical diffusivity on height and degree of turbulence are known to agree reasonably well with experiments in the lower atmosphere, and are used by O G. Sutton (1934) and W G L. Sutton (1943), in their work on evaporation problems They are

$$\frac{U}{U_1} = \left(\frac{z}{z_1}\right)^m, \tag{4 1}$$

and

$$K_z = a_z U_1^{1-n} z^{1-m}, \tag{4·2}$$

where $m = n/(2-n)$ is a non-dimensional parameter, taken as a measure of the degree of turbulence.

In these expressions U_1 is the value of the mean wind speed at some standard height z_1, a_z is a constant for a given turbulent state, involving n, z_1 and the physical constants of the fluid; U_1, in the z_1 plane, is assumed to be independent of x and y, so that U and K_z are functions of z only. The coefficient a_z used is that given by O G Sutton (1934)

$$a_z = \left\{\frac{(\frac{1}{2}\pi k_z^2)^{1-n}(2-n)^{1-n}n^{1-n}}{(1-n)(2-2n)^{2-2n}}\right\}\lambda^n z_1^{(n^2-n)/(2-n)}, \tag{4 3}$$

where λ has been identified with the kinematic viscosity of air by O G Sutton (1934), and k_z is the Karman constant for wind-tunnel conditions In the open atmosphere the scale magnification in eddy effect is allowed for and an 'atmospheric k_z' calculated

Some assumption must now be made with regard to K_y, and a reasonable one appears to make K_y obey the wind-power law Comparison with equation (4·2) suggests

$$K_y = a_y U_1^{1-n} z^n. \tag{4·4}$$

By analogy with equation (4·3) and keeping K_y dimensionally correct, it is assumed that

$$a_y = \left\{\frac{(\frac{1}{2}\pi k_y^2)^{1-n}(2-n)^{1-n}n^{1-n}}{(1-n)(2-2n)^{2-2n}}\right\}\lambda^n z_1^{(n^2-4n+2)/(2-n)} \tag{4·5}$$

5. Wind speed and diffusivities variable with height above the evaporating surface

The initial step in the solution of this problem is obtained by writing

$$\xi = x, \quad \zeta = \frac{2}{(2m+1)}\left(\frac{U_1^n}{a_z z_1^m}\right)^{\frac{1}{2}} z^{m+\frac{1}{2}},$$

and

$$\eta = \left(\frac{U_1^n}{a_y z_1^m}\right)^{\frac{1}{2}} y. \tag{5·1}$$

The transport equation (2·2) then becomes in these co-ordinates, after insertion of K_y and K_z in accordance with equations (4·2) and (4·4),

$$\zeta^{1-2p}\frac{\partial \chi}{\partial \xi} = \frac{\partial}{\partial \zeta}\left[\zeta^{1-2p}\frac{\partial \chi}{\partial \zeta}\right] + \frac{\partial}{\partial \eta}\left[\zeta^{1-2p}\frac{\partial \chi}{\partial \eta}\right], \tag{5·2}$$

where

$$p = \frac{n}{(2+n)}.$$

The original bounding parabola $y^2 = 4Ax$ becomes $\eta^2 = 4\mu\xi$, where now

$$\mu = \frac{U_1^n}{a_y z_1^m} A$$

With

$$X = \xi, \quad Y = \frac{\eta}{2\sqrt{(\mu\xi)}}, \quad Z = \frac{\zeta}{2\sqrt{(\mu\xi)}},$$

equation (5·2) becomes

$$X\frac{\partial \chi}{\partial X} = \tfrac{1}{2}Y\frac{\partial \chi}{\partial Y} + \tfrac{1}{2}Z\frac{\partial \chi}{\partial Z} + \frac{1}{4\mu}\left[\frac{\partial^2}{\partial Y^2} + \frac{\partial^2}{\partial Z^2}\right]\chi + \frac{(1-2p)}{4\mu Z}\frac{\partial \chi}{\partial Z}. \tag{5·3}$$

As in the parallel case of constant wind and diffusivity a solution independent of X exists and has then to satisfy

$$\frac{\partial^2 \chi}{\partial Z^2} + \frac{\partial^2 \chi}{\partial Y^2} + 2\mu\left[Y\frac{\partial \chi}{\partial Y} + Z\frac{\partial \chi}{\partial Z}\right] + \frac{(1-2p)}{4\mu Z}\frac{\partial \chi}{\partial Z} = 0. \tag{5·4}$$

Substituting $Y + iZ = c\cosh(\beta + i\theta)$, equation (5·4) becomes

$$\frac{\partial^2 \chi}{\partial \beta^2} + \frac{\partial^2 \chi}{\partial \theta^2} + 2\mu\cosh\beta\sinh\beta\frac{\partial \chi}{\partial \beta} - 2\mu\cos\theta\sin\theta\frac{\partial \chi}{\partial \theta}$$
$$+ (1-2p)\coth\beta\frac{\partial \chi}{\partial \beta} + (1-2p)\cot\theta\frac{\partial \chi}{\partial \theta} = 0. \tag{5·5}$$

A solution independent of θ exists, and one which is consistent with the boundary conditions is

$$\frac{\chi}{\chi_0} = 1 - \frac{\int_0^\beta \exp[-\mu\sinh^2\phi](\sinh\phi)^{2p-1}d\phi}{S(\mu, p)}, \tag{5·6}$$

where

$$S(\mu, p) = \int_0^\infty F(\mu, p, \phi)\,d\phi$$
$$= \int_0^\infty \exp[-\mu\sinh^2\phi](\sinh\phi)^{2p-1}d\phi. \tag{5·7}$$

It appears that for the values of p obtained in the atmosphere ($0 < p < \frac{1}{3}$), this integral is not expressible in terms of known functions

The rate of evaporation at a point (x, y) on the saturated area is given by equation (3·9), and the substitutions (3·3) and (5·1), to be

$$E(x,y) = \frac{\chi_0 a_z U_1^{1-n}(m+\frac{1}{2})^{1-2p}(U_1^n/z_1^m a_z)^p}{S(\mu,p)(\sin\theta)^{2p}(2\sqrt{[\mu x]})^{2p}}. \tag{5·8}$$

Then, since $\sin\theta = \frac{1}{2}\sqrt{[(4Ax-y^2)/Ax]}$,

$$E(x,y) = c/(4Ax-y^2)^p, \tag{5·9}$$

where

$$C = \chi_0 \frac{a_z U_1^{1-n}(m+\frac{1}{2})^{1-2p}}{S(\mu,p)}\left(\frac{a_y}{a_z}\right)^p. \tag{5·10}$$

The total rate of evaporation from a parabolic area of downwind length x is given by

$$E = 2C\int_0^x\int_0^{2\sqrt{(Ax)}} \frac{1}{(4Ax-y^2)^p}\,dx\,dy.$$

By substituting $y = 2\sqrt{(Ax)}\sin t$, this becomes

$$E = 2C(4A)^{\frac{1}{2}-p}\int_0^x x^{\frac{1}{2}-p}\,dx\int_0^{\frac{1}{2}\pi}(\cos t)^{1-2p}\,dt, \tag{5·11}$$

and hence

$$E = \chi_0 \frac{2^{2-2p}a_z U_1^{1-n}(m+\frac{1}{2})^{1-2p}(a_y/a_z)^p A^{\frac{1}{2}-p}\,\Gamma(\frac{1}{2})\,\Gamma(1-p)\,x^{\frac{1}{2}-p}}{S(\mu,p)(3-2p)\,\Gamma(\frac{1}{2}[3-2p])}. \tag{5·12}$$

This function reduces, as is to be expected, to the equation (3·12) when p is equated to the value $\frac{1}{2}$

It should be noted that this solution is rigorously true only for parabolic areas infinite in extent downwind, with vertices pointing upwind and axes lying parallel to the mean wind direction. For a parabolic shape, of finite length x_1, lying upwind of the line $x = x_1$ the formula (5·9) is not applicable in the very small region near to this bounding edge, but it is reasonable to suppose that the total rate of evaporation given by equation (5·12) is not affected to any significant extent In the same way equation (5·6) for the concentration χ is only true as far as $x = x_1$.

Equation (5·9) enables the relation between β and χ/χ_0 to be evaluated and hence the iso-concentration lines in any vertical plane at $x = x_1$ are then given by

$$\frac{y^2}{4Ax\cosh^2\beta} + \frac{z^{2m+1}}{4(m+\frac{1}{2})^2 Ax_1\sinh^2\beta(a_y/a_z)} = 1. \tag{5·13}$$

6 Wind speed and normal diffusivity variable with height and lateral diffusivity zero

For purposes of comparison it is necessary to obtain the theoretical equations giving the concentration of vapour above and evaporation from a parabolic area, on the assumption of wind speed and normal diffusivity varying with height and lateral diffusivity zero.

When $$\frac{U}{U_1} = \left(\frac{z}{z_1}\right)^m, \quad K_x = a_x U_1^{1-n} z^{1-m} \quad \text{and} \quad K_y = 0,$$

the well-established solution due to O. G. Sutton (1934) for evaporation from a rectangle of downwind length x_0 and width dy is

$$E = K U_1^{(2-n)/(2+n)} x_0^{2/(2+n)} dy, \tag{6·1}$$

where $$K = \chi_0 \left(\frac{2+n}{2-n}\right)^{(2+n)/(2-n)} \left(\frac{2+n}{2\pi}\right) \sin\frac{2\pi}{(2+n)} \Gamma\left(\frac{2}{2+n}\right) a_z^{2/(2+n)} z_1^{-n^2/(4-n^2)} \tag{6·2}$$

Under these conditions the solution for the parabola $y^2 = 4Ax$ becomes

$$E(x_1) = 2K U_1^{(2-n)/(2+n)} \int_0^{2\sqrt{(Ax)}} \left[x_1 - \frac{y^2}{4A}\right]^{2/(2+n)} dy$$

$$= 2K U_1^{(2-n)/(2+n)} x_1^{(6+n)/(4+2n)} \sqrt{(A\pi)} \frac{\Gamma\left(\dfrac{4+n}{2+n}\right)}{\Gamma\left(\dfrac{10+3n}{4+4n}\right)}. \tag{6·3}$$

The theoretical distribution of vapour concentration over a saturated area of infinite lateral extent (i e. in effect, neglecting lateral dispersion) was obtained by O. G. Sutton (unpublished) This gives the concentration χ at any distance x from the upwind edge and height z above ground-level, in terms of the ground concentration χ_0 in the form

$$\frac{\chi}{\chi_0} = 1 - I\left[\frac{U_1^n z^{(2+n)/(2-n)}}{\left(\dfrac{2+n}{2-n}\right)^2 a_z x z_1^{n/(2-n)}}, \frac{-2}{(2+n)}\right] \tag{6·4}$$

$I(r, p)$ is tabulated in Pearson's *Tables of the Incomplete Γ-Functions*.

On this assumption of zero lateral dispersion the solution for the parabola $y^2 = 4Ax$, of downwind length x, along the downwind edge becomes

$$\frac{\chi}{\chi_0} = 1 - I\left[\frac{U_1^n z^{(2+n)/(2-n)}}{\left(\dfrac{2+n}{2-n}\right)^2 a_z \left(x_1 - \dfrac{y^2}{4A}\right) z_1^{n/(2-n)}}, \frac{-2}{(2+n)}\right] \tag{6·5}$$

7 Comparison of theoretical results with experiment

The solution in § 3, based on the assumption of constant wind and diffusivities, is of interest only from the theoretical point of view in that it indicates the mathematical approach to the harder case where these entities are allowed to vary with height. No numerical results are given.

The main difficulty in the theoretical computation of equations (5·6) and (5·12) lies in evaluating the integral $S(\mu, p)$. Simpson's rule was found to be sufficiently accurate for comparison with experiment, but difficulty was encountered near the origin where the integral tended to an infinite value This was overcome by sub-

stituting the asymptotic series for the sinh function and integrating over a small range 0 to ϵ, with $p = \frac{1}{5}$,

$$\int_0^{\epsilon} F(\mu, \phi)\, d\phi = \tfrac{5}{2}\epsilon^{2/5} - \tfrac{5}{20}(\mu + \tfrac{7}{54})\epsilon^{20/5}, \qquad (7{\cdot}1)$$

where $F(\mu, \phi)$ is given in equation (5·7). The integral from ϵ to infinity was evaluated by Simpson's rule and added to the integral from 0 to ϵ, giving the complete $S(\mu, p)$

Experiments were carried out some time ago by the author in a Ministry of Supply wind tunnel to measure rates of evaporation from parabolic surfaces, saturated with aniline, of equal area, arranged in turn on the floor of the working section of the tunnel with axes parallel to the mean air stream direction and vertices directed upstream. The experimental technique followed that used by Pasquill in experiments described fully in his paper (Pasquill 1943); and evaporation, air temperature, and wind speed in the free stream of the working section were measured as in his experiments. The wind-velocity 'profile' (i.e. variation of wind with height) near the evaporating plate was determined by means of a small, standard, pitot static tube. these observations yielded a measure of the degree of turbulence given by the index m of equation (4·1) of 0·120 in good agreement with previous values. The observed rates of evaporation were reduced to the form E/χ_0 g./min /unit of saturated vapour concentration at the surface (c g s. units), where E is the total rate of evaporation for the given area and χ_0 is the saturation vapour concentration at the temperature of the air stream. E/χ_0 was chosen as it is independent of the temperature.

Some typical theoretical and measured results for wind-tunnel scale of turbulence are given in table 1.

TABLE 1. CALCULATED AND MEASURED RATES OF EVAPORATION, E/χ_0, EXPRESSED AS GRAMMES PER MINUTE PER UNIT (C.G S) OF SATURATED VAPOUR CONCENTRATION AT THE SURFACE, FOR PLANE PARABOLIC SHAPES OF AREA = 208 SQ.CM , IN WIND-TUNNEL CONDITIONS

A (cm.)	x (cm)	type of result	wind speed in m./sec. at 1 cm. height above surface: 2	4	6
0 50	23·0	*calculated* with $K_y \propto z^m$; $K_z \propto z^{1-m}$	9,600	16,600	23,100
0 50	23·0	*calculated* with $K_y = 0$; $K_z \propto z^{1-m}$	9,180	16,000	22,200
0·50	23 0	*measured*	11,000	19,600	27,600
3 00	12·7	*calculated* with $K_y \propto z^m$; $K_z \propto z^{1-m}$	9,920	17,200	23,950
3 00	12 7	*calculated* with $K_y = 0$, $K_z \propto z^{1-m}$	9,600	16,800	23,000
3·00	12·7	*measured*	11,700	21,500	30,700

This table shows that, for this set of experiments, the '$K_y = 0$ theory' underestimates the rates of evaporation by 13 to 20 %, and that the use of K_y, as specified in § 4, improves agreement by 2 to 4%. In this connexion, a much more comprehensive series of experiments by Pasquill (1943) on evaporation from circular and rectangular areas in a wind tunnel gave even better agreement (to within 5 % for wind speeds

up to 5 m./sec.) between '$K_y = 0$ theory' and experiment. It has therefore been established that for practical purposes the theoretical model of turbulence with one-dimensional dispersion is already satisfactory for assessments of *evaporation*.

The important difference in the two theories is expected in the decay of concentration in liquid vapour with height above the saturated surface. Unfortunately, reliable experimental results for rates of variation above parabolic surfaces are not available, but it is of interest to *compare* theoretical values Table 2 was calculated by applying the equations for concentration to an area 10 yd downwind length and bounded by the curve $y^2 = 8x$. A so-called 'atmospheric k' was used for estimating a_y and a_z, when $n = 0{\cdot}25$ and wind speed (at 2 m.) = 4 m./sec : i e the von Kármán k for wind-tunnel flow was multiplied by a factor, dependent on meteorological observations, to allow for the scale differences of eddy effect in the wind tunnel and in the open atmosphere. Use was made of ideas due to Prandtl and Th. von Kármán. These are described by O. G Sutton (1934). If $|\omega'|$ denote the absolute magnitude of vertical component of eddy velocity, then

$$|\omega'| = l'\left|\frac{\partial \bar{u}}{\partial z}\right|,$$

where l' is an individual 'mixing length' Hence $\overline{|\omega'|} = l\left|\frac{\partial \bar{u}}{\partial z}\right|$; l is the mean mixing length. This equation, together with von Kármán's expression

$$l = k\left|\frac{\partial \bar{u}}{\partial z}\right| \Big/ \left|\frac{\partial^2 \bar{u}}{\partial z^2}\right|,$$

yields the equation

$$\overline{|\omega'|} = k\left|\frac{\partial \bar{u}}{\partial z}\right|^2 \Big/ \left|\frac{\partial^2 \bar{u}}{\partial z^2}\right|.$$

The appropriate values of k were determined from (a) equation (4·1) and (b) experimental values of $\overline{|\omega'|}$.

TABLE 2. THEORETICAL DECAY IN CONCENTRATION ABOVE THE CENTRE OF THE DOWNWIND EDGE OF A PARABOLIC AREA, $y^2 = 8x$, OF DOWNWIND LENGTH 10 YARDS IN THE OPEN ATMOSPHERE

U (at 2 m) = 4 m /sec
$n = 0\ 25$ (i e. approximately zero temperature gradient conditions)

Ratio of $\dfrac{\text{concentration of vapour at height } Z \text{ in}}{\text{concentration of vapour at height 1 in.}}$ above the centre of the downwind edge.

theory	$K_y = 0$				$K_y = a_y u_1^{1-n} z^m$			
height (in.)	1	8	24	66	1	8	24	66
ratio	1·00	0·43	0 14	0 011	1·00	0·29	0·06	0 005

Experiments have been carried out by the Ministry of Supply to determine the decay with height in concentration of vapour above rectangular areas (some typical results are shown in table 3). It was discovered that the calculated decay, using

the one-dimensional dispersion system, is much too slow. These experimental values are grouped closely around zero temperature gradient conditions and have been averaged. Theoretical, i.e. '$K_y = 0$', results based on these average conditions, have been worked out from equations (6·4) and are quoted in table 3 They are identical with the '$K_y = 0$' results of table 2 (same 'n' value, with slightly different wind speed) to the order of accuracy quoted. Table 2 illustrates the much more rapid dilution of concentration obtained with the two-dimensional system.

TABLE 3. MEASURED AND CALCULATED CONCENTRATION OVER THE DOWNWIND EDGE OF A RECTANGULAR AREA 10 YARDS DOWNWIND LENGTH AND 30 YARDS LATERAL DIMENSIONS SAMPLING POINT AT THE CENTRE OF THE DOWNWIND EDGE

wind speed in m /sec (at 2 m. height)	temperature gradient °F (56′ − 3′)	ratio of concentration at height Z in / concentration at height 1 in			
		1 in	8 in.	24 in	66 in.
		(*a*) measured values			
4·10	− 1 7	1 00	0 24	0 06	0 002
3·21	− 1·7	1 00	0 39	0 06	0 003
4·19	0	1 00	0·33	0 08	0 007
5 84	− 0 1	1 00	0 28	0 04	0 003
5·56	− 0 4	1·00	0 35	0 05	0·014
means 4·58	− 0 6	1 00	0 31	0 06	0 005
		(*b*) calculated values (based on $K_y = 0$)			
4 58	0	1·00	0 43	0·14	0 011

8 DISCUSSION

It appears from § 7 and other results (Pasquill 1943; Frost 1946) that for plane, finite evaporating surfaces the theoretical expressions for total rates of evaporation into turbulent air streams, on the basis of an eddy diffusion agency operative only in a direction normal to the evaporating surface, is adequate for practical purposes A slight improvement is introduced with the expressions obtained in § 5, but the advantage in simplicity of mathematical form lies with previous results The equations obtained by the one-dimensional model, however, for the decay in concentration above an evaporating strip of *finite* lateral dimensions seriously underestimate the effect It is proposed that the introduction of a lateral diffusion mechanism, even though empirical in nature, is a first step towards an improved theory Solution of the resulting differential equation for areas of other shapes has as yet not been successful, e g. a rectangular strip, infinite in extent downwind, appears to lead to Mathieu functions of the hyperbolic and elliptic type in the case of constant wind and diffusivity, and to the Associated Mathieu functions in the general case A completely satisfactory analytical solution has, as yet, defied all attacks. The question of a second lateral edge has also not been resolved.

It will be of interest to specify atmospheric diffusivity by means of Frost's empirical formula (Frost 1946) and to compare results for evaporation from a parabolic area with those obtained using O. G Sutton's expressions. These calculations are being carried out by the author.

For the future development of the theory of turbulence in the atmosphere there is a clear need for a new formulation which will furnish precise expressions for the coefficients of eddy diffusion. At the moment it appears that a good approximation only has been obtained by postulating the behaviour of these coefficients in the atmosphere.

I wish to express my thanks to Dr W. G. Bickley for his initial suggestion and guidance, and to Professor L. Rosenhead, F R S , for his interest in and criticism of the work. Finally, acknowledgement is made to the Controller General of Research and Development, Ministry of Supply, and to the University of London for permission to publish this paper. The contents of table 3 are taken from a Ministry of Supply Report, and a part of this paper was included as a section of the author's thesis for the degree of Ph D of the University of London

References

Frost, R. 1946 *Proc Roy Soc.* A, **186**, 20
Pasquill, F 1943 *Proc Roy. Soc* A, **182**, 75
Sutton, O. G 1932 *Proc. Roy Soc.* A, **135**, 668
Sutton, O G. 1934 *Proc. Roy Soc.* A, **146**, 701
Sutton, W. G. L 1943 *Proc Roy. Soc.* A, **182**, 48

The vibration spectrum and molecular configuration of cyclohexane

By D. A Ramsay and G B B M Sutherland

Laboratory of Colloid Science, Cambridge

(*Communicated by E. K. Rideal, F.R.S.—Received 2 September* 1946)

The physical evidence for the structure of the cyclohexane molecule (X ray, electron diffraction, dipole moment, vibration spectra) is all in favour of the 'chair' configuration. The twelve vibration frequencies for the carbon skeleton of this model have been calculated, using a simple valency force field There are eight distinct frequencies, four being doubly degenerate The two force constants are evaluated from two of the Raman frequencies and used to calculate the remaining six frequencies Five of these frequencies are found to agree with the observed frequencies to within 4 %, while the sixth frequency is too low to be observed conveniently in the infra-red and is forbidden in the Raman spectrum. The value calculated (206 cm.$^{-1}$) for this sixth frequency is, however, consistent with the specific heat data. The selection rules are shown to be obeyed by the assigned frequencies The value for the C-C stretching force constant, viz. $3\ 7 \times 10^5$ dynes/cm is considerably lower than that for the same bond in ethane, viz. $4\ 5 \times 10^5$ dynes/cm , indicating a weakening of the C-C bond in cyclohexane relative to ethane

Introduction

It is more than fifty years since Sachse (1890) first pointed out that it was possible for the cyclohexane molecule to exist in two configurations—the well-known 'chair' and 'boat' forms—without distorting the tetrahedral valencies of the carbon atoms. Later, in their work on the stereochemistry of certain cyclohexane derivatives, Perkin & Pope (1906) adopted a planar configuration for the cyclohexane ring, though this involved considerable distortion of the carbon valence angles (figure 1)

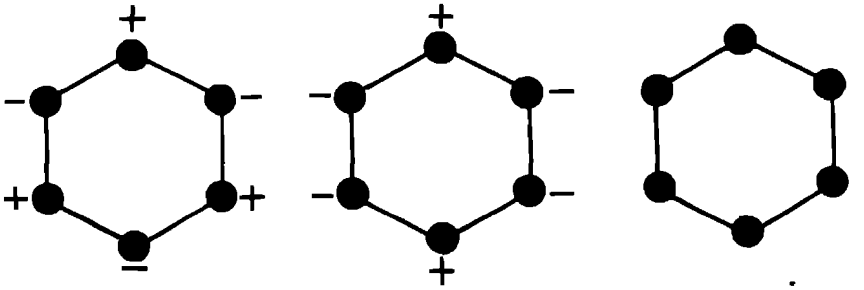


Figure 1

In recent years several physical methods have been applied to the determination of the configuration of the cyclohexane ring and considerable evidence has accumulated in favour of the 'chair' structure. Hassel & Kringstad (1930) investigated the crystal structure of cyclohexane at − 80° C and showed that the results favoured the

'chair' configuration. Similar investigations by Dickinson & Bilicke (1928) on the crystal structure of hexachloro- and hexabromo-cyclohexane, and by Halmöy & Hassel (1932) on the crystal structure of 1 : 4-diiodo- and 1 . 4-dibromo-cyclohexane support this view Lonsdale & Smith (1939), however, found that an X-ray powder photograph of cyclohexane at $-180°$ C indicated the probable existence of a low-temperature modification of lower symmetry. Pauling & Brockway (1937) showed that the electron diffraction pattern for cyclohexane vapour was inconsistent with a planar configuration for the molecule, but their results did not distinguish conclusively between the 'chair' and 'boat' forms, though the theoretical intensity curves for the 'chair' form were in better agreement with the experimental curves Similar conclusions were reached by Hassel & Taarland (1940) Extensive measurements on the dipole moments of cyclohexane derivatives have been carried out by Hassel (1934). In particular the dibromo- and diiodo-derivatives of high melting point were found to have zero dipole moments, indicating that the cyclohexane ring exists in a highly symmetrical form. Similar results were found for the corresponding dichloro-compound and for the hexachloro-derivative.

The different symmetry properties of the three structures enables one to distinguish between them by the different numbers of fundamental vibration frequencies active in the infra-red and Raman spectra in the three cases and further by differences in the states of polarization of the Raman lines. Langseth & Bak (1940) investigated the Raman spectrum of cyclohexane and concluded that their polarization data supported a planar configuration, but Kohlrausch & Wittek (1941) found an error in the polarization measurements and showed that the revised polarization data were consistent only with the 'chair' configuration. These measurements were subsequently confirmed by Gerding, Smit & Westrik (1942). The infra-red spectrum of cyclohexane vapour has been examined by Rasmussen (1943) and the number of active fundamentals found to be compatible with the 'chair' configuration only

In this paper we have calculated the vibration frequencies of the 'chair' structure using a simple valency force field, to determine whether a quantitative interpretation of the experimental frequencies can be obtained and to derive information concerning the strengths of the C-C bonds in the molecule.

Equations of motion

The cyclohexane molecule is treated as a system of six equal mass points, i, j, k, l, m, n, each of mass m, situated at the vertices of a puckered hexagon whose geometry is defined so that the angles γ between adjacent sides are all tetrahedral, and the bond lengths s are all equal (figure 2)

Applying a valency force field to the molecule, the potential energy V may be expressed in the form

$$V = \tfrac{1}{2}\sum_{ij} k q_{ij}^2 + \tfrac{1}{2}\sum_{i} c\alpha_{jk}^{i2},$$

where k = stretching constant of the C-C bond, c = deformation constant of the C-C-C angle, q_{ij} = change in the bond length s_{ij}, α_{jk}^{i} = change in the bond angle γ_{jk}^{i}.

Employing the vector method recently developed by Eliashevich (1940) the following equations are obtained for the motion of the system.

$$\left.\begin{aligned}
q_{ij} &= -\frac{k}{m}(2q_{ij}+\cos\gamma\,\overline{q_{ik}+q_{jl}})+\frac{c\sin\gamma}{2ms}(2\alpha^i_{jk}+2\alpha^j_{il}-\alpha^l_{jn}-\alpha^k_{im}),\\
\alpha^i_{jk} &= \frac{k\sin\gamma}{2ms}(2q_{ij}+2q_{ik}-q_{jl}-q_{km})\\
&\quad -\frac{c}{ms^2}(4\alpha^i_{jk}-\alpha^j_{il}-\alpha^k_{im}+\alpha^l_{jn}+\alpha^m_{kn})\\
&\quad +\frac{c\cos\gamma}{2ms^2}(4\alpha^i_{jk}-2\alpha^j_{il}-2\alpha^k_{im}-\alpha^l_{jn}-\alpha^m_{kn})
\end{aligned}\right\} \qquad (1)$$

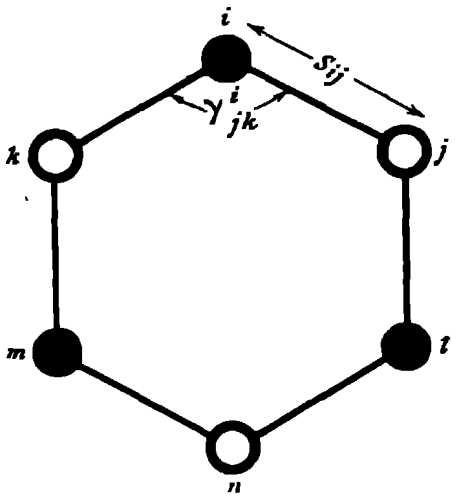


FIGURE 2. ● Mass above the plane of the paper, ○ mass below the plane of the paper

Assume that the system executes simple harmonic oscillations and substitute

$$q_{ij} = (q_{ij})_0\, e^{i\nu t} \quad \text{and} \quad \alpha^i_{jk} = (\alpha^i_{jk})_0\, e^{i\nu t}$$

equations in ν^2 are obtained whose roots determine the frequencies of the normal modes of vibration of the system. These equations are considerably simplified if we utilize the symmetry properties of the molecule.

SYMMETRY PROPERTIES

The symmetry elements of the 'chair' structure for the cyclohexane molecule consist of (*a*) a centre of inversion, (*b*) a principal threefold axis, (*c*) three equivalent twofold axes perpendicular to the principal axis, (*d*) three equivalent planes of symmetry intersecting in the principal axis, and are usually represented by the group symbol D_{3d} (figure 3).

This group consists of the following classes of symmetry operations:

(*a*) The identity operation (I) which leaves the molecule unchanged.

(*b*) Inversion about the centre (i)

(*c*) Rotation of $\pm\frac{2\pi}{3}$ about the threefold axis ($2C_3$).

(*d*) Rotation of π about the twofold axes ($3C_2$).

(*e*) Rotation of $\pm\frac{2\pi}{6}$ about the threefold axis, followed by reflexion in a plane perpendicular to this axis ($2S_6$).

(*f*) Reflexion across the planes of symmetry ($3\sigma_d$).

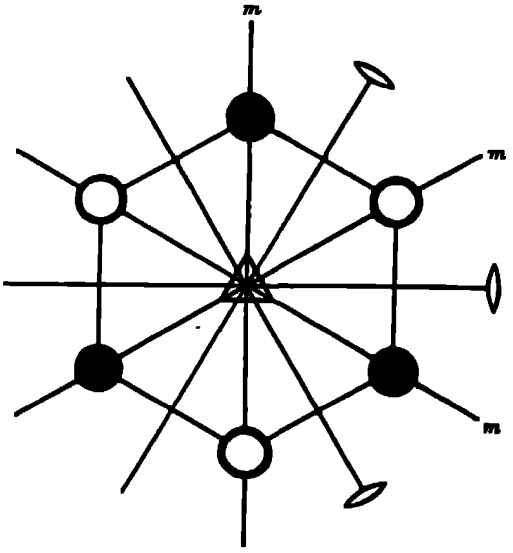


FIGURE 3. m, plane of symmetry, (), twofold axis

The symmetry properties of the various classes of vibration in this group are given in table 1.

TABLE 1

class	I	$2C_3$	$3C_2$	i	$2S_6$	$3\sigma_d$
A_{1g}	+	+	+	+	+	+
A_{1u}	+	+	+	−	−	−
A_{2g}	+	+	−	+	+	−
A_{2u}	+	+	−	−	−	+
E_{1g}	+	ϵ^{120}	±	+	$-\epsilon^{60}$	±
E_{1u}	+	ϵ^{120}	±	−	$+\epsilon^{60}$	±

The symbols A_1 and A_2 refer to non-degenerate classes of vibration, while E_1 denotes classes of vibration which are doubly degenerate The subscripts g and u denote classes of vibration which are respectively symmetrical or antisymmetrical with respect to the centre of inversion. Non-degenerate vibrations must either be

symmetrical (+) or antisymmetrical (−) with respect to each of the symmetry operations, i e the displacements of the nuclei after the operation of a symmetry element must either remain unchanged or be reversed in direction. Doubly degenerate vibrations may be symmetrical or antisymmetrical with respect to certain symmetry operations, but are transformed in a special manner by the remaining symmetry elements The two components of a degenerate pair are mutually orthogonal, i e

$$\sum_i m_i \xi_{ia} \xi_{ib} = 0,$$

where m_i = mass of the ith particle, ξ_{ia}, ξ_{ib} = displacement vectors of the ith particle in the component vibrations, and may be selected in an infinite number of ways. The degenerate pair will be chosen so that one component is symmetrical and the other antisymmetrical with respect to one of the planes of symmetry This is denoted in table 1 by the symbol $\pm$. The two components then behave in a similar manner with respect to the twofold axis at right angles to this plane of symmetry With respect to rotation about the threefold axis, the two components transform according to the equations

$$\xi'_{ia} = \xi_{ia} \cos 120° + \xi_{ib} \sin 120°,$$

$$\xi'_{ib} = -\xi_{ia} \sin 120° + \xi_{ib} \cos 120°,$$

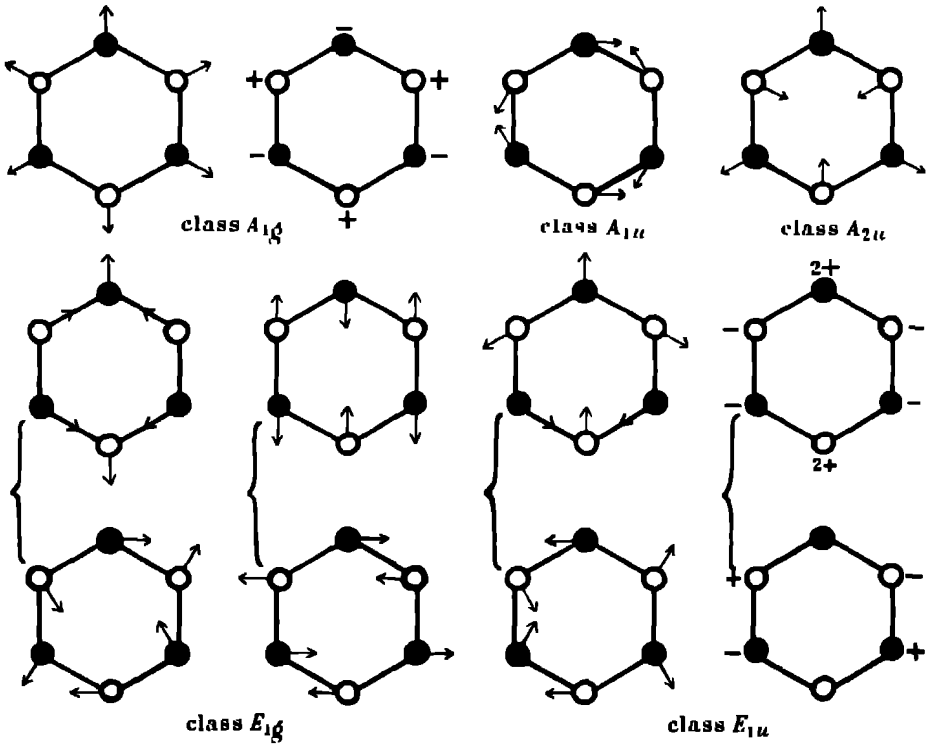


FIGURE 4

where ξ'_{ia}, ξ'_{ib} denote the displacement vectors after rotation. This is denoted by the symbol ϵ^{120} The symmetry operations $2S_6$ are similarly represented by the transformation $\pm\epsilon^{60}$.

The geometrical forms of the normal vibrations may now be derived using the above symmetry properties and the principle that every pair of normal vibrations must be mutually orthogonal. These are given in figure 4. There are no ring vibrations in the class A_{2g}

Expressing these vibrations in terms of q the change of bond length and α the change in the C-C-C angle, the sets of symmetry co-ordinates given in figure 5 are obtained.

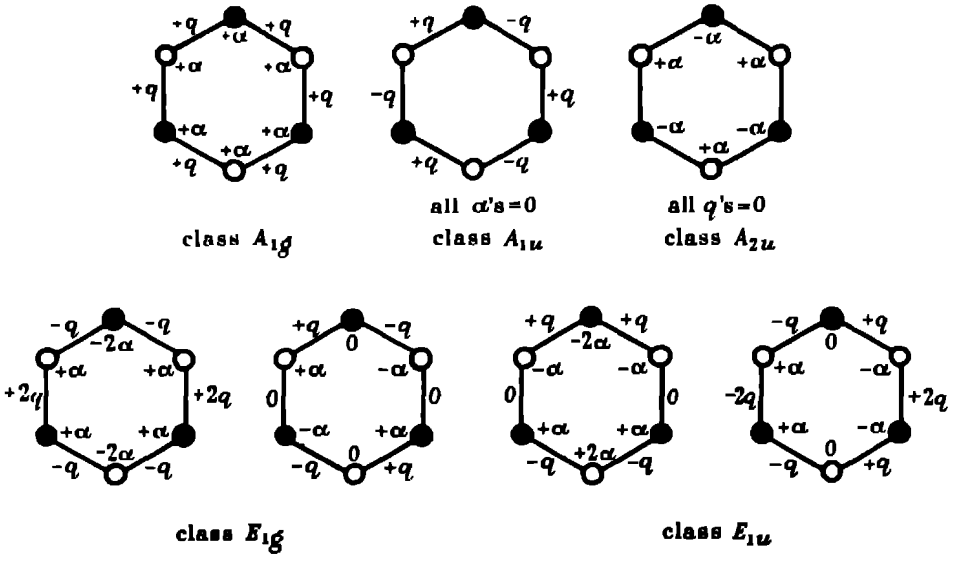


FIGURE 5

VIBRATION EQUATIONS

By substituting the appropriate symmetry co-ordinates into the equations for q_{ij} and α^i_{jk} (equations (1)), the following equations for the various classes of vibration are obtained.

$$\text{Class } A_{1g} \qquad \nu_1^2+\nu_2^2=\frac{2k}{m}(1+\cos\gamma)+\frac{c}{ms^2}(4+\cos\gamma),$$

$$\nu_1^2\nu_2^2=\frac{4ck}{m^2s^2},$$

$$\text{Class } A_{1u} \qquad \nu^2=\frac{2k}{m}(1-\cos\gamma);$$

$$\text{Class } A_{2u}. \qquad \nu^2=\frac{9c}{ms^2};$$

$$\text{Class } E_{1g}. \quad \nu_1^2+\nu_2^2 = \frac{k}{m}(2-\cos\gamma)+\frac{c}{2ms^2}(8-7\cos\gamma),$$

$$\nu_1^2\nu_2^2 = \frac{17ck}{2m^2s^2};$$

$$\text{Class } E_{1u}. \quad \nu_1^2+\nu_2^2 = \frac{k}{m}(2+\cos\gamma)+\frac{c}{2ms^2}(4-3\cos\gamma),$$

$$\nu_1^2\nu_2^2 = \frac{3ck}{2m^2s^2}.$$

Experimental

We have investigated the infra-red absorption spectrum of cyclohexane vapour in the region 1500 to 500 cm.$^{-1}$ using a Hilger D 209 double beam spectrometer with rock salt and potassium bromide prisms. The cyclohexane used was a specially pure sample obtained from I.C.I., Billingham. The length of the absorption path

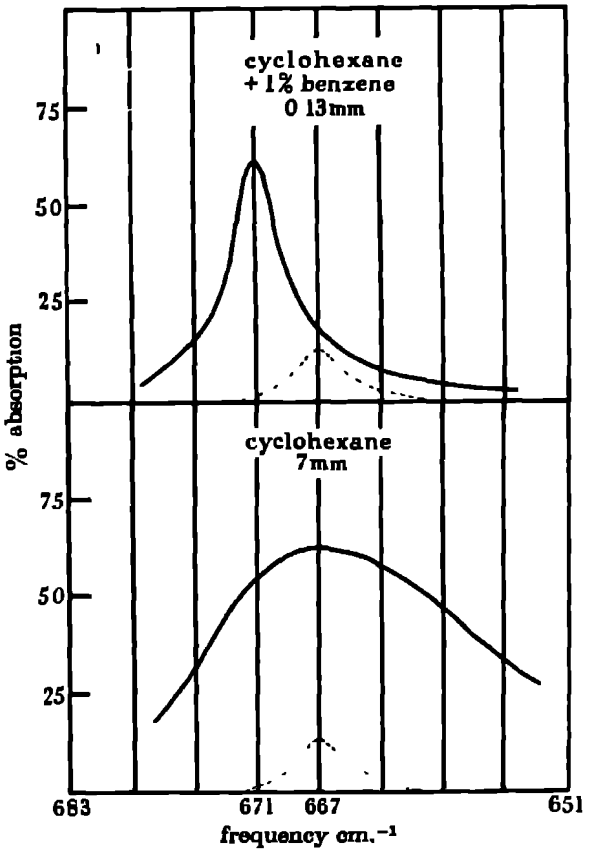


Figure 6. --- Atmospheric CO_2 absorption used for calibration.

was 25 cm and the vapour pressure of cyclohexane 88 mm. The spectrum obtained was in good agreement with the earlier curves published by Kettering & Sleator (1933) and Rasmussen (1943), except that we did not observe the absorption band at 14·85μ (673·4 cm.$^{-1}$) recorded by the former authors This band is possibly due to a small amount of benzene impurity, as it agrees in position and contour with a strong benzene band recorded by Kettering & Sleator (1933) We have also measured the infra-red absorption of cyclohexane in the liquid state using an absorption path of 0·05 mm. The principal frequencies agreed with those observed in the vapour, but additional weak bands were observed at 1350, 1158, 1110, and 524 cm $^{-1}$. Again we observed no band at 673 4 cm.$^{-1}$. By increasing the absorption path to 7 mm , we did, however, observe a weak band in this region at 667 cm $^{-1}$ That this was not due to a small amount of benzene impurity was shown by comparison with a spectrum of cyclohexane to which 1 % benzene has been added (figure 6) It is certain that the position and shape of the cyclohexane absorption differs markedly from that of benzene in this region

No measurements have been made on the Raman spectrum of cyclohexane vapour, but the Raman spectrum of the liquid has been investigated by Langseth & Bak (1940), Kohlrausch & Wittek (1941) and Gerding, Smit & Westrik (1942) The infra-red and Raman frequencies below 1500 cm.$^{-1}$ are listed in table 2, the figures in brackets giving the relative intensities of the bands

TABLE 2. INFRA-RED AND RAMAN FREQUENCIES FOR CYCLOHEXANE

infra-red		Raman
vapour	liquid	liquid
—	524 (2)	384 (1), *P*
—	667 (0)	426 (2), *D*
864 (6)	862 (6)	802 (10), *P*
903 (5)	905 (6)	1029 (6), *D*
1015 (1)	1017 (3)	1157 (3), *P*
1042 (3)	1041 (4)	1267 (6), *D*
—	1110 (0)	1348 (2), *D*
—	1158 (0)	1426 (½), —
1261 (5)	1259 (6)	1444 (6), *D*
—	1350 (0)	1465 (½), —
1456 (10)	1453 (10)	—

P, polarized; *D*, depolarized.

ASSIGNMENT OF FUNDAMENTALS

The forty-eight normal modes of vibration of cyclohexane may, to a first approximation, be divided into the following groups. 12 C-H stretching vibrations, 6 H-C-H bending vibrations, 6 C-C stretching vibrations, 6 C-C-C bending vibrations and 18 $>CH_2$ rocking and twisting vibrations. The symmetry classes of these vibrations, assuming a 'chair' configuration with symmetry D_{3d}, have been discussed by

Rasmussen (1943) and are given in table 3. The frequency ranges for the various types of vibration and the infra-red and Raman selection rules are also given It should be noted that no frequency is permitted to appear both in the infra-red and the Raman spectrum as this model has a centre of symmetry.

TABLE 3. SYMMETRY CLASSES OF THE VARIOUS TYPES OF VIBRATION

type of vibration	frequency range (cm $^{-1}$)	number of vibrations per class					
		A_{1g}	A_{1u}	A_{2g}	A_{2u}	E_{1g}	E_{1u}
12 C-H stretching	2850–3000	2	0	0	2	2	2
6 H-C-H bending	1350–1500	1	0	0	1	1	1
18 $>CH_2$ rocking and twisting	700–1350	1	2	2	1	3	3
6 C-C stretching	700–1350	1	1	0	0	1	1
6 C-C-C bending	<700	1	0	0	1	1	1
selection rules		*P*	*f*	*f*	*IR*	*D*	*IR*

P, polarized line in Raman, *f*, forbidden;
D, depolarized line in Raman, *IR*, active in infra-red

The three polarized lines in the Raman spectrum below 1500 cm.$^{-1}$, viz 384, 802, 1157 cm.$^{-1}$, may be assigned to the totally symmetrical class of vibrations A_{1g} The 384 cm.$^{-1}$ line is undoubtedly the symmetrical C-C-C bending frequency, and the 802 cm $^{-1}$ line is probably the symmetrical C-C stretching frequency (cf the symmetrical breathing frequency of benzene at 992 cm.$^{-1}$). Substituting these frequencies into the equations for the class A_{1g} vibrations the following values for the force constants are obtained.

$$k = 3{\cdot}69 \times 10^5 \text{ dynes/cm.}, \quad c/s^2 = 0{\cdot}438 \times 10^5 \text{ dynes/cm}$$

Using these force constants the frequencies for the remaining classes of vibration may be calculated. These are given in table 4.

TABLE 4. OBSERVED AND CALCULATED FREQUENCIES FOR CYCLOHEXANE

class	activity	cm.$^{-1}$ (calc.)	cm $^{-1}$ (obs.)	% diff.
A_{1g}	Raman	802	802	—
		384	384	—
A_{1u}	inactive	1092	1110	−1 6
A_{2u}	infra-red	691	667?	+3 6
E_{1g}	Raman	1068	1029	+3·7
		420	426	−1 4
E_{1u}	infra-red	914	903	+1·2
		206	?	—

The weak frequency at 1110 cm $^{-1}$ observed in the infra-red spectrum of the liquid but absent in the spectrum of the vapour is probably the inactive class A_{1u} fundamental appearing in violation of the selection rules It is interesting to note

that two of the other additional weak bands observed in the spectrum of the liquid, viz 1158 and 1350 cm $^{-1}$, coincide with frequencies permitted in the Raman spectrum. A breakdown of the selection rules in the liquid state is well known.

The depolarized lines in the Raman spectrum at 1020 and 426 cm.$^{-1}$ lie close to the frequencies calculated for the class E_{1g} vibrations. Confirmation of the degeneracies of these frequencies is afforded by the Raman spectrum of cyclohexane—d_1 investigated by Langseth & Bak (1940). Each of these frequencies is found to be split into two components, due to the reduction of the symmetry of the molecule by the introduction of the deuterium atom

The infra-red frequency at 903 cm $^{-1}$ agrees closely with the frequency evaluated for the class E_{1u} stretching vibration, but direct confirmation of the degenerate bending frequency in this class has not been possible owing to the experimental difficulty of measuring infra-red frequencies lower than 300 cm.$^{-1}$ A value of 210 cm.$^{-1}$ for this frequency has been derived from specific heat measurements by Aston, Szasz & Fink (1943), but the agreement with our calculated value of 206 cm.$^{-1}$ cannot be regarded as confirmatory evidence owing to the drastic assumptions regarding the magnitudes of the 18 $>CH_2$ rocking frequencies made by these authors However, it is significant that in the closely related molecules tetrahydropyran, 1 4-dioxane and trioxane, a Raman frequency has been observed in each case by Kahovec & Kohlrausch (1937) in the region 200 to 250 cm $^{-1}$

The assignment for the class A_{2u} frequency is less certain We have been unable to detect any absorption in the region of 691 cm $^{-1}$ in the infra-red spectrum of the vapour or in the spectrum of the liquid using an absorption path of 0 05 mm , but there is always the possibility that this fundamental is exceptionally weak in absorption We therefore investigated the absorption of a 7 mm. layer of the liquid in this region and found a weak band at 667 cm $^{-1}$ We have shown (figure 6) that this band is not due to benzene impurity and hence may be the weak fundamental required.

This completes the assignment of the skeletal frequencies which we have computed, and it is not profitable to attempt a fuller interpretation of the cyclohexane spectrum This cannot be done until the spectrum has been investigated using very high resolving power, and due consideration has been given to the theory of the hydrogenic frequencies However, it should be possible to interpret all frequencies below 700 cm $^{-1}$, and the only one which we have not yet accounted for is the weak infra-red band at 524 cm $^{-1}$. This is most probably the difference band (903–384) cm $^{-1}$.

We have also considered whether the available evidence on band contours is consistent with our assignments. The 'chair' form of cyclohexane is a symmetrical top molecule Assuming values of 1·09 and 1·54 A respectively for the C-H and C-C interatomic distances, the moments of inertia are

$$I_A = I_B = 194 \times 10^{-40}\,\text{g cm.}^2, \quad I_C = 335 \times 10^{-40}\,\text{g.cm.}^2$$

According to the formulae of Gerhard & Dennison (1933), this means that the parallel bands should have a P-R doublet separation of 22·5 cm.$^{-1}$ at 15° C. The perpen-

dicular bands should also have a P, Q, R structure in which the P-R doublet separation is decidedly smaller. It is not possible to give this separation precisely from their curves, and in addition it must be remembered that Gerhard & Dennison (1933) took no account of Coriolis's interaction in the perpendicular bands, so that their calculations are only qualitatively correct for many perpendicular bands The only infra-red band in our assignments which has been observed in the vapour state is that at 903 cm $^{-1}$. This should be a perpendicular band and so would not be expected to show a doublet separation of 22 5 cm.$^{-1}$ This is consistent with the existing experimental data.

The values of the force constants

The value for the stretching constant of the C-C bond in cyclohexane obtained above, viz. $3{\cdot}7 \times 10^5$ dynes/cm , is seen by reference to table 5 to be considerably lower than any of the values obtained by various authors for ethane. Differences in force constants of bonds in polyatomic molecules must be interpreted with caution, as the bond-force constant may vary appreciably with the potential function employed. However, we believe the difference observed here to be a real one and to indicate a weakening of the C-C bond in cyclohexane compared with ethane, since (*a*) the difference is much greater than the variations in the ethane constant found by using different functions, and (*b*) the function we have used contains no interaction terms (which usually alter bond constants appreciably) and yet gives a very satisfactory interpretation of the skeletal spectrum Dr A. D Walsh has pointed out to us that additional confirmation of our conclusion is provided by the fact that the ionization potential of cyclohexane is considerably less than that of ethane (Walsh 1946)

Table 5

author	substance	C-C stretching constant $\times 10^5$ dynes/cm.	C-C-C bending constant $\times 10^5$ dynes/cm
Stitt (1939)	ethane	4 50	—
Linnett (1940)	ethane	4·53	—
Eliashevich &	ethane	4·47	—
Stepanov (1941)	propane	4·47	0 70
	n-butane	4·47	0 70
Present	cyclohexane	3·7	0 44

It is interesting to note that the value for the deformation constant of the bond angle is very much lower than that given by Eliashevich & Stepanov (1941) for propane and *n*-butane. Although the assumption made by these authors that the C-H and C-C force constants in propane and *n*-butane are the same as those in ethane is not entirely justified, the difference between our value and theirs is too great to be explained as other than a real effect. It would appear, therefore, that the decrease in the C-C bond stretching force constant on cyclization of a saturated hydrocarbon

is accompanied by a corresponding decrease in the C-C-C bond angle deformation constant.

We are grateful to Professor E. K. Rideal for his interest and encouragement, to Dr R P. Bell, Dr A. D. Walsh and Mr T. E W. Nind for valuable discussion and to Messrs I. C. I., Billingham, for the sample of pure cyclohexane. One of us (D.A R) wishes to acknowledge a maintenance grant from Messrs B. X. Plastics Ltd., during the tenure of which this work was carried out.

References

Aston, J. G., Szasz, G. J & Fink, H L 1943 *J. Amer Chem. Soc.* **65**, 1135.
Dickinson, R G & Bilicke, C 1928 *J Amer. Chem. Soc* **50**, 764.
Eliashevich, M. A. 1940 *C R. Acad Sci U R S.S* **28**, 604.
Eliashevich, M A. & Stepanov, B. 1941 *C R. Acad. Sci. U.R S.S.* **32**, 481.
Gerding, H , Smit, E & Westrik, R 1942 *Rec. trav. chim. Pays-Bas*, **61**, 561.
Gerhard, S L. & Dennison, D. M 1933 *Phys. Rev* **43**, 197.
Halmöy, E. & Hassel, O. 1932 *Z. Phys. Chem.* B, **16**, 234.
Hassel, O. 1934 *Trans Faraday Soc.* **30**, 874
Hassel, O. & Kringstad, H. 1930 *Tidsskr. Kemi*, **10**, 128.
Hassel, O & Taarland, T. 1940 *Tidsskr. Kemi*, **20**, 167
Kahovec, L. & Kohlrausch, K. W. F. 1937 *Z. Phys. Chem.* B, **35**, 20.
Kettering, C. F. & Sleator, W. W. 1933 *Physics*, **4**, 39
Kohlrausch, K. W. F & Wittek, H 1941 *Z. Phys. Chem.* B, **48**, 177
Langseth, A & Bak, B 1940 *J Chem. Phys.* **8**, 403
Linnett, J. W 1940 *J. Chem. Phys* **8**, 91.
Lonsdale, K. & Smith, H 1939 *Phil Mag.* **28**, 614.
Pauling, L & Brockway, L. O 1937 *J. Amer. Chem. Soc* **59**, 1223.
Perkin, W. H. & Pope, W. J. 1906 *Proc Chem Soc.* **22**, 107
Rasmussen, R. S. 1943 *J. Chem. Phys* **11**, 249
Sachse, H. 1890 *Ber dtsch. chem. Ges.* **23**, 1363.
Stitt, F. 1939 *J. Chem. Phys.* **7**, 297
Walsh, A. D. 1946 *Trans. Faraday Soc* **42**, 779.

The crystal structure of sucrose sodium bromide dihydrate

By C. A. Beevers and W. Cochran

Dewar Crystallographic Laboratory, University of Edinburgh

(*Communicated by Sir Lawrence Bragg, F.R.S.—Received 4 October 1946*)

The complete crystal structure of sucrose sodium bromide dihydrate has been determined. The crystals are orthorhombic, with space-group $P2_12_12_1$, and cell-edges $a = 21.92$, $b = 9.72$, $c = 8.43$ A. The bromine positions were obtained from Patterson-Harker syntheses, and the signs of the F's of the $(0kl)$, $(h0l)$ and $(hk0)$ planes from the change of intensity from the chloride to bromide compounds. Fourier projections parallel to the three axes were used to determine the complete structure.

In the structure the pyranose ring of the molecule is of the Sachse *trans* form. The furanose ring is an unexpectedly compact grouping, and one member of the ring is displaced from the plane of the other four. The configuration of both rings is such as to allow the groups attached to the ring atoms to approach as nearly as possible to the mean plane of each ring. Direct proof that sucrose may be described as

α-*d*-glucopyranosido-β-*d*-fructofuranoside

has been obtained for the first time. The sucrose molecules are held together by bonds from the hydroxyl groups to the Na and Br ions and to the water molecules, the hydroxyl groups each having two external bonds.

The method used of reducing the observed intensities to an absolute scale is described, and the proof of the correctness of the structure is given as a set of diagrams of calculated and observed electron density.

Introduction

The crystal structure of sucrose sodium bromide dihydrate is of interest for the following reasons. (1) It forms a valuable exercise in the solution of a complex crystal. The sugars are well known to be difficult subjects for complete X-ray analysis. The difficulties arise partly from a lack of knowledge of the detailed conformation of the saccharide molecules, and a few accurate determinations should give considerable assistance in subsequent work on other sugars. (2) This particular structure is that of an addition compound of an unusual type containing different types of chemical bonds, the external bonding of the —OH groups of the sugar molecule being of especial interest. (3) The sucrose molecule provides evidence of the detailed geometrical form of both the pyranose and furanose rings, about which there has been considerable speculation. The pyranose ring was supposed to have all its carbon atoms in a plane by Cox, Goodwin & Wagstaff (1935), who argued from the cell dimensions of over sixty sugars and their methyl derivatives. Later, however, the structure of glucosamine hydrobromide (and of the isomorphous hydrochloride) was determined by Cox & Jeffrey (1939), who found the pyranose ring to be of the Sachse *trans* form. Hitherto there has been no evidence as to the conformation of the furanose ring in sugars, but it has often been supposed to be flat, since this gives interbond angles of 108° within the ring, a good approximation to the tetrahedral angle. (4) The chemical evidence for the linkage of the pyranose and furanose rings

through an oxygen atom is conclusive, but it is not certain from chemical results whether the monose residues have the α or the β configuration. Indirect evidence favours the α form for the glucopyranoside residue and the β form for the fructofuranoside residue, so that the formula for sucrose should probably be as shown in figure 1, but X-ray structural work appears to be the only method capable of giving a clear decision on these points

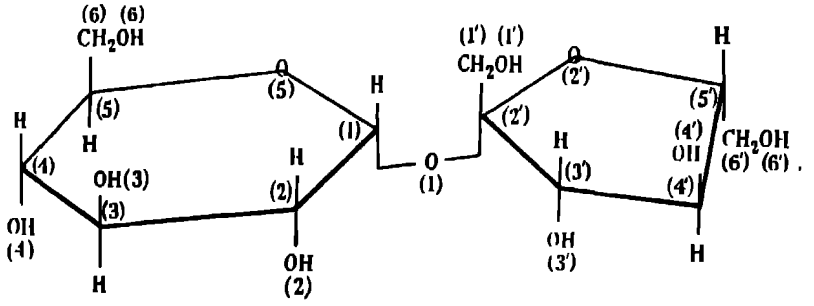


FIGURE 1 Chemical structure of sucrose, showing most probable configurations of C(1) and C(2').

A brief description of certain addition compounds containing sucrose has been given by one of us in a previous communication (Cochran 1946) The substances investigated were

$2C_{12}H_{22}O_{11}$ $3NaI$ $3H_2O$, $C_{12}H_{22}O_{11}.NaBr$ $2H_2O$ and $C_{12}H_{22}O_{11}$ $NaCl$ $2H_2O$

The second and third of these compounds are isomorphous, and the projection of their space-group $P2_12_12_1$ down a crystal axis gives the plane group Pba, which possesses centres of symmetry These facts make it possible to overcome the fundamental difficulty of X-ray analysis—that the relative phases of the diffracted beams are unknown—and determine by purely synthetic methods the crystal structure of $C_{12}H_{22}O_{11}$ $NaBr$ $2H_2O$, provided only that the co-ordinates of the replaceable atom can be found

EXPERIMENTAL

Specimens for photography were prepared by chipping away parts from a single crystal with a razor blade until a roughly cylindrical portion remained, the length of the cylinder being in the direction of an axis of the crystal The cylinder was then smoothed by rolling it between two pieces of fine sandpaper until its diameter was 1·1 mm Weissenberg photographs of individual layer lines were taken with the crystal oscillating through an angle of 240°. About twelve such layer lines result from an a-axis specimen when Cu $K\alpha$ radiation is used; these were photographed as far as the ninth. Four layer lines were photographed from each of the b- and c-axis specimens. Three photographs of exposure time 10 hr., 2 hr , and $\frac{1}{2}$ hr. were taken of each. The complete investigation of the two compounds thus involved the prepara-

tion and orientation of six crystal specimens and the taking of over one hundred Weissenberg photographs. So far only the data obtained from the zero layer-line photographs have been utilized to determine the crystal structure. The following table summarizes the results obtained by measurements made on these photographs and in other ways

	a	b	c	V	z	density (calc.)	density (obs.)	μ	space-group
$C_{12}H_{21}O_{11}$ NaBr $2H_2O$	21·92	9·72	8·43	1796	4	1·778	1·783	43·4	$P2_12_12_1$
$C_{12}H_{21}O_{11}$ NaCl $2H_2O$	21·75	9·62	8·40	1758	4	1·651	1·650	29·4	$P2_12_12_1$

Densities were found experimentally by adding alcohol to bromoform until a crystal fragment remained suspended in the mixture. The space-group was found from the systematic absences $h00$ for h odd, $0k0$ for k odd, and $00l$ for l odd, no other systematic absences being observed.

Intensity estimation

The assignment of intensities to the photographically recorded X-ray reflexions was done by a method which we believe to be more accurate but no less rapid than that previously employed in this laboratory. The previous method consisted in the visual comparison of spots on the photograph with a photographic wedge in the centre of the film, and it suffers from two main disadvantages. (1) The wedge consists of twenty steps, each of uniform photographic density, while the spots vary in density across their breadth because of absorption of the X-ray beam in the crystal. Efficient visual comparison of uniform and non-uniform densities is difficult. (2) The spots are viewed against a background of lesser photographic density, while the scale is usually imprinted on a part of the photograph that has not been exposed at all. Thus for accuracy two observations should be made, one of the intensity of the spot and one of the background intensity in its immediate neighbourhood, the latter should then be subtracted from the former. In attempting to estimate intensities by this method it was noticed that it was much easier to judge the departure from equality of two spots close together on the film than it was to compare either of them with the wedge in the centre. This observation forms the basis of the method finally adopted. By exposing different parts of a film to the same X-ray reflexions for various lengths of time a number of spots whose photographic densities corresponded to X-ray intensities in a known ratio were obtained. Intensities could now be estimated by moving this scale beside the spot under examination until one was found on the scale which appeared identical when viewed by diffusely transmitted light. The two spots were thus viewed through the same background and, since their densities varied in roughly the same way, comparison was facilitated. By using the photographs of different exposure times which had been obtained for each layer line, intensities ranging from 1 to 400 were placed on a numerical scale. The accuracy of this very simple method can be assessed from the following table, which shows

the intensities assigned to the ($h00$) reflexions given by both b- and c-axis specimens of $C_{12}H_{22}O_{11}$ NaBr $2H_2O$

index h	2	4	6	8	10	12	14	16	18	20	22	24	26	28
b-axis	0	224	144	220	224	3	6	42	0	30	0	$2\frac{1}{2}$	$5\frac{1}{2}$	30
c axis	0	220	150	255	250	4	8	44	0	31	0	2	$5\frac{1}{2}$	31

Polarization and Lorentz factors

The calculation of these effects was made in the following way. Zero layer sections of the reciprocal lattice were drawn out on a suitable scale, and the intensity $I(hk)$ of each reflexion was written beside the corresponding lattice point (hk) Each such point is at $2\sin\theta$ from the lattice origin Values of $\Theta = (1+\cos^2 2\theta)/(\sin 2\theta)$, at a suitable interval, were written out on a piece of transparent material, each value of Θ being written at a distance of $2\sin\theta$ from a pin-hole near one end of the transparent scale By pivoting this scale about a pin passing through the origin of the lattice section the value of Θ corresponding to each lattice point could be read off directly This method could be used to find rapidly any factor, for each reflexion, which is a function of $\sin\theta$. The value of $10I/\Theta = G^2$ was now calculated for each reflexion, the factor 10 being introduced to make the numbers G of a convenient order of magnitude

Location of replaceable atom

The bromine atom was located by Patterson-Harker methods. As is well known the Patterson function exhibits maxima at vector distances from the origin equal to vector distances between pairs of atoms.

$$P(uv) = \frac{1}{A^2}\Sigma\Sigma\, F^2(hk0)\cos 2\pi(hu+kv)$$

represents this function projected on (001), its maxima give the components of the vectors in this plane The equivalent points of $P2_12_12_1$ are·

$$(xyz)\ (\tfrac{1}{2}+x, \tfrac{1}{2}-y, \bar{z})\ (\bar{x}, \tfrac{1}{2}+y, \tfrac{1}{2}-z)\ (\tfrac{1}{2}-x, \bar{y}, \tfrac{1}{2}+z).$$

Vectors between these points have components u and v shown below·

u	$\frac{1}{2}$	$\frac{1}{2}$	$2x$	$\overline{2x}$	$\frac{1}{2}+2x$	$\frac{1}{2}-2x$	$\frac{1}{2}+2x$	$\frac{1}{2}-2x$
v	$\frac{1}{2}+2y$	$\frac{1}{2}-2y$	$\frac{1}{2}$	$\frac{1}{2}$	$2y$	$\overline{2y}$	$\overline{2y}$	$2y$

Vectors of this type occur for each atom in the unique volume of the unit cell, and $P(uv)$ will exhibit corresponding maxima In particular $P(u\frac{1}{2})$ will have a large peak at $u = \pm 2x_1$, while $P(\frac{1}{2}v)$ will have a large peak at $v = \frac{1}{2} \pm 2y_1$, where x_1 and y_1 are co-ordinates of the bromine atom. Because of the symmetry of the projection these functions may be written (omitting proportionality constants)

$$P(u\tfrac{1}{2}) = \sum_h \{\Sigma\,(-1)^k F^2(hk0)\}\cos 2\pi hu$$

and

$$P(\tfrac{1}{2}v) = \sum_k \{\Sigma\,(-1)^h F^2(hk0)\}\cos 2\pi kv.$$

Similar considerations apply to the (yz) and (xz) co-ordinates, so that two independent values can be got for each of x_1, y_1 and z_1. Values of F^2 in absolute units were not available at this stage of the investigation, values of G^2 (as defined above) were used instead as coefficients in the one-dimensional syntheses. This method gave

$$x_1 = 0{\cdot}065,\ 0\ 065, \quad y_1 = 0\ 175,\ 0{\cdot}178; \quad z_1 = 0\ 048,\ 0{\cdot}053$$

Isomorphous replacement method

The theory of what may be called the isomorphous replacement method of crystal-structure determination will now be briefly outlined, taking the projection on (001) as an example of its application. This projection has centres of symmetry at the points $(\frac{1}{4}0)\,(\frac{3}{4}0)\,(\frac{1}{4}\frac{1}{2})\,(\frac{3}{4}\frac{1}{2})$. If the origin is transferred to $(\frac{1}{4}0)$ it will therefore coincide with a centre of symmetry. The atomic structure factor referred to this origin is

$$S(hk) = \quad 4\cos 2\pi hx \cos 2\pi ky \quad \text{when} \quad h+k = 2n$$

and

$$S(hk) = -4\sin 2\pi hx \sin 2\pi ky \quad \text{when} \quad h+k = 2n+1.$$

Assuming that all atoms have the same co-ordinates in the crystal structures of $C_{12}H_{22}O_{11}.NaBr.2H_2O$ and $C_{12}H_{22}O_{11}.NaCl\ 2H_2O$ (denoted by subscripts B and C respectively), then

$$F_B = Sf_B + \sum_i S_i f_i \quad \text{and} \quad F_C = Sf_C + \sum_i S_i f_i,$$

where f_B, f_C and f_i are atomic scattering factors and subscript i denotes atoms other than bromine or chlorine.

Therefore
$$F_B - F_C = S(f_B - f_C). \tag{1}$$

The right-hand side of this equation involves quantities which can be evaluated from the known position of the bromine atom. If $|F_B|$ and $|F_C|$ are known, the signs to be associated with each can therefore be determined. It may be noted that in order to make use of this result each $|F|$ must be known in absolute units, whereas the method of intensity measurement outlined above gives $|G|$, which represents $|F|$ on an arbitrary scale and is uncorrected for a number of effects. Write

$$F^2 \propto G^2ATR.$$

Now these functions, the absorption factor A, the temperature effect T, and the factor R which has been introduced to take account of the fact that the photographic method does not measure integrated reflexion intensity but peak reflexion intensity, are all functions of $\sin\theta$, at least for zero layer-line reflexions. Hence write $F = gG$, where g is a correction factor which varies with $\sin\theta$. To establish the structure amplitudes on an absolute scale it is therefore necessary to find $g_B(\sin\theta)$ and $g_C(\sin\theta)$. This can be done by making use of equation (1) in the form

$$g_B G_B - g_C G_C = S(f_B - f_C). \tag{2}$$

All observations were grouped together in about a dozen sets, all those of one set corresponding to approximately the same value of $\sin\theta$ The values of g_B and g_C could therefore be regarded as constant within each group Approximate values of these factors were obtained by inspection of a number of cases where one of G_B or G_C happened to be zero. These approximate values, when inserted in equation (2) were sufficient to establish the signs to be associated with $|G_B|$ and $|G_C|$ in the great majority of cases. More accurate values were then obtained by setting up as many equations of type (2) as possible, involving G's of approximately the same value of $\sin\theta$ and then solving by the method of least squares for g_B and g_C. These are shown as functions of $\sin\theta$ in the accompanying graph (figure 2) Each G was now multiplied by the corresponding value of g to give F Calculated and observed values of $F_B - F_C$ could now be compared and the agreement was found to be sufficiently good to establish the signs of the F's in the great majority of cases. A number of discrepancies remained as long as it was assumed that the chlorine atom had exactly the same co-ordinates as the bromine, but these could almost all be removed by assuming that the x co-ordinate of chlorine was 0·058, instead of 0 065 as found for bromine This was confirmed by evaluating the function $P_C(u\frac{1}{2})$ which was found to have a maximum corresponding to $x = 0\ 0575$, but did not have one corresponding to $x = 0\ 065$ By applying similar tests to the y and z co-ordinates these were found to be unchanged in passing from bromine to chlorine. By using equation (1) in the modified form

$$F_B - F_C = S_B f_B - S_C f_C$$

it was now possible to establish the signs of a sufficient number of the F's (about 80 %) to carry out the computation of electron density maps.

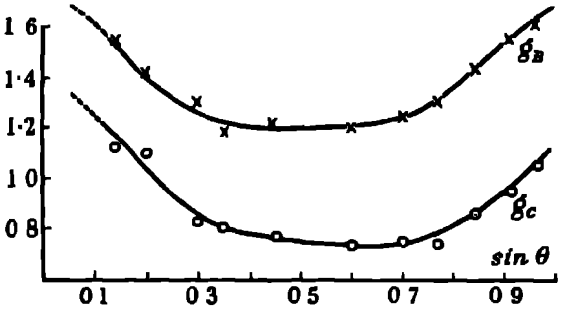


FIGURE 2 Graph of the values of g_B and g_C.

INTERPRETATION OF FOURIER PROJECTIONS

The structure was by no means obvious from the Fourier projections since the omission of important terms from the syntheses caused certain atomic positions to be represented by very diffuse maxima, while other maxima were later found not to correspond to atomic positions The sodium atom was readily located since it caused electron density maxima considerably greater than those due to carbon or

oxygen, and it was known that its distance from the bromine atom would be approximately 3 A The sucrose structure was arrived at in the following way Copies of the electron density maps on a scale of 1 in. = 1 A were supported by a wooden framework along the three sides of the unit cell Three electric lamps were arranged at a distance of several metres so that each gave an almost parallel beam of light, perpendicular to one of the planes A number of small spheres with holes bored to accommodate tetrahedrally disposed bonds of steel rod 1 5 in long were used to build up a rough molecular model, starting from a single atom and adding others in such a way that the shadows cast always fell on maxima of all three electron density projections In this way a satisfactory structure was arrived at for one-half of the molecule—the glucose residue Considerable difficulty was experienced in settling the approximate configuration of the fructose residue since it does not give a really clear projection on any of the three planes. A number of arrangements were tried until one was found which gave very satisfactory agreement between calculated and observed structure amplitudes, and it was then possible to find by calculation the signs of all important terms previously omitted from the syntheses The electron density maps were then recalculated and found to be in very good agreement with the postulated structure (see later) Final atomic co-ordinates were derived almost entirely from the (010) and (001) projections since practically none of the atoms is separately resolved in the (100) projection due to the comparatively small area of the latter The eighty-one parameters which define the structure are given below (in 1/60ths of the cell edges)

Parameters in 60ths of cell edges

Br	(3 0	10 6	3·1)	$CH_2(1')$	(15 0	56 0	38 0)
Na	(4 4	28 3	7·5)	OH(1')	(15 8	48 3	33 8)
CH(1)	(11 2	12 8	36 0)	C(2')	(10 9	58 0	38 0)
O(1)	(10 3	6 3	42 0)	O(2')	(9·5	57·0	28 6)
CH(2)	(13 2	19 2	41·5)	CH(3')	(8 5	52·8	43 4)
OH(2)	(16 2	16·6	45 8)	OH(3')	(8 4	54 0	53·5)
CH(3)	(10 4	23 5	47 6)	CH(4')	(5 2	51 8	37 8)
OH(3)	(12 3	30 2	51 8)	OH(4')	(3 0	44 5	40 8)
CH(4)	(7 2	26 5	43 0)	CH(5')	(6 5	51 4	28 1)
OH(4)	(4 4	30 0	49 0)	$CH_2(6')$	(4 0	54·1	21 1)
CH(5)	(5 4	19 0	37·3)	OH(6')	(1 8	0·8	22 9)
O(5)	(8 1	15 8	30 2)				
$CH_2(6)$	(2 0	22 0	31 0)	$H_2O(1)$	(10 6	23·5	14 0)
OH(6)	(3 3	28 0	24 5)	$H_2O(2)$	(7 9	40 7	5 0)

Description of structure

Figure 3 gives a projection of the complete structure on the a-b plane The two-fold screws parallel to the c-axis pass through the points $x = \frac{1}{4}$ and $\frac{3}{4}$, $y = 0$ and $\frac{1}{2}$ on this diagram The screws parallel to the b-axis pass through the points $x = 0$ and $\frac{1}{2}$, $z = \frac{1}{4}$ and $\frac{3}{4}$, while the screws parallel to a pass through the points $y = \frac{1}{4}$ and $\frac{3}{4}$, $z = 0$ and $\frac{1}{2}$. These symmetry elements operate on the atomic positions given in the

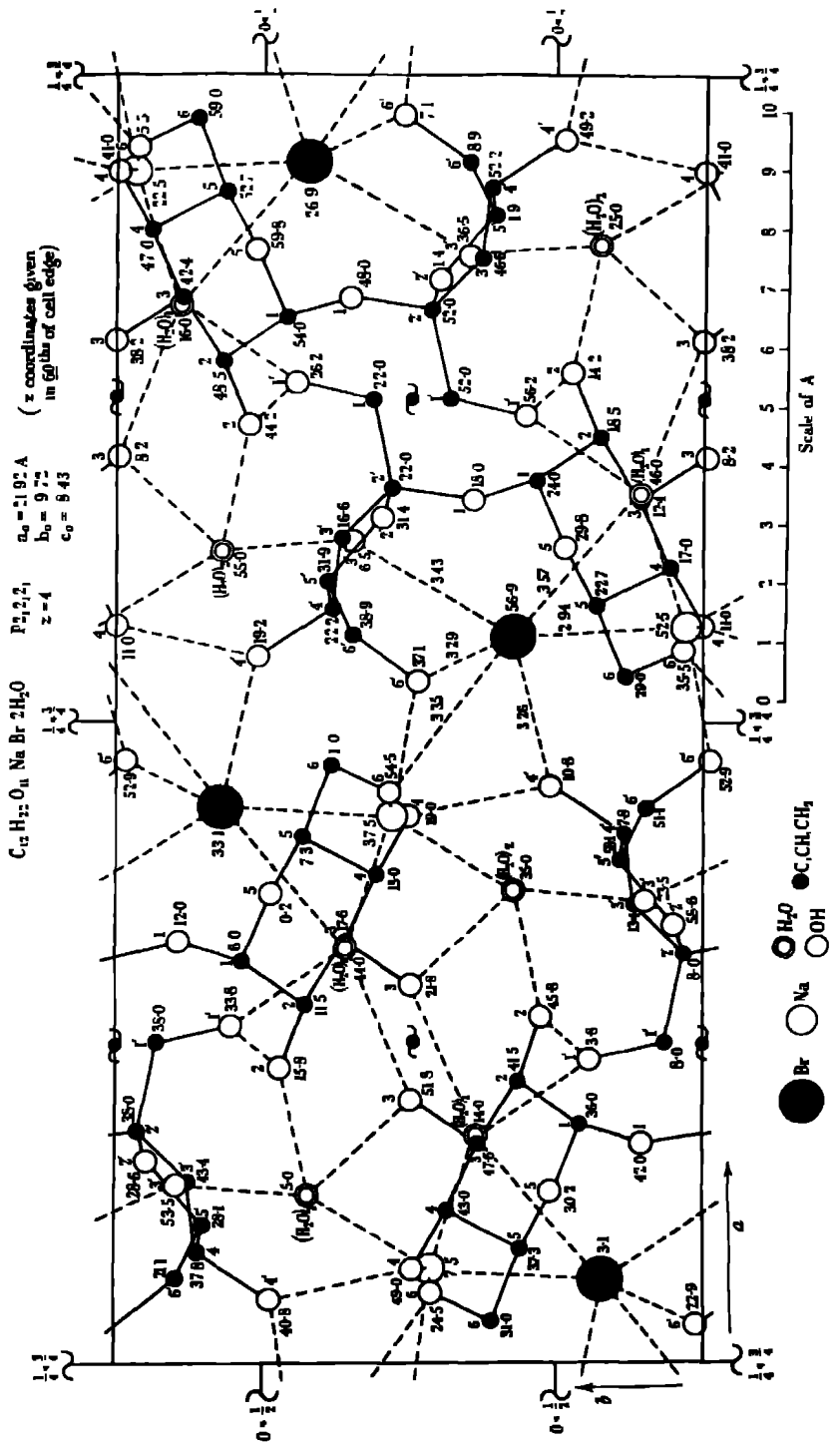


FIGURE 3. Projection of the complete structure on to (001).

last paragraph and in this way produce the complete structure Each atom or molecule in figure 3 is labelled in accordance with the numbering of figure 1, and its z co-ordinate in 60ths of the z axis is written alongside.

Consider first of all the structure of the sucrose molecule itself. Figure 4 shows a perspective drawing of this molecule taken out of the complete structure. In this diagram the bond lengths are inserted It is clear from this diagram that the oxygen atoms attached to carbons 1 and 2 of the pyranose ring {(O)1 and (OH)2} are in the *cis* relation to one another, so that the glucose residue possesses the alpha configuration. Also the oxygens attached to carbons 2′ and 3′ of the fructose ring are in the *cis* relation to one another, so that the fructose residue possesses the beta configuration This establishes with certainty that sucrose may be described as

α-*d*-glucopyranosido-β-*d*-fructofuranoside.

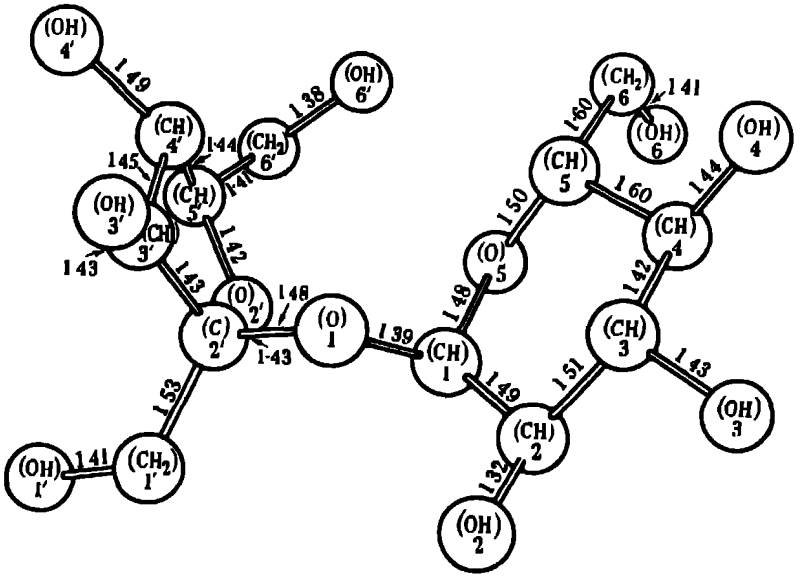


FIGURE 4. Perspective drawing of sucrose molecule.

The pyranose ring is of the Sachse *trans* form with the following interatomic distances.

	$C_{12}H_{22}O_{11}.NaBr\ 2H_2O$	α chitosamine hydrobromide
C1–C2	1·48$_5$A	1·48A
C2–C3	1·51	1 58
C3–C4	1·42	1 56
C4–C5	1 60	1·55
C5–O5	1 50	1 50
O5–C1	1 48	1·37

This ring is very similar indeed to the ring in α chitosamine hydrobromide (Cox & Jeffrey 1939) except that Cox & Jeffrey have chosen the mirror-image molecule.

The distances between the atoms of the pyranose ring in α chitosamine are also given above Some of the distances agree well in the two structures, whilst others are different probably by more than experimental error The average C-C distance within the ring is 1 50 A (in α chitosamine hydrobromide it is 1·54 A), and this is perhaps not significantly different from the single bond value 1 54 A The bond angles of the carbons in the ring are also not appreciably different from the tetrahedral angle, their mean value is 108° and the mean deviation is only 4°, perhaps about equal to the probable error The *trans* form of the ring is the form which allows the atoms of the ring together with their immediate addenda (except for the linking oxygen O 1) to be flattened as much as possible while keeping the tetrahedral angles.

The furanose ring possesses some unexpected features The five members of the ring are definitely not in one plane Four of them, however, namely C2′, C3′, C5′ and O2′ are rather accurately coplanar, but the other member of the ring C4′ is away from this plane by about ½ A Imagine this ring to be produced from a completely plane ring by a rotation of the portion C3′ C4′ C5′ about the line C3′ C5′. This rotation of C3′ and C5′ is such as to bring the side groups to which they are bonded ((OH)3′ and (CH_2)6′ respectively) more nearly into the plane C2′ C3′ C5′ O2′ and the rotation and translation of C4′ are together sufficient to bring its attached group (OH)4′ almost exactly into the plane C2′ C3′ C5′ O2′ Thus it seems that the departure from planarity of the ring itself is once more in the interest of a more general planarity of the residue as a whole The conformation is very nearly that which would be obtained if the ring C2′ C3′ C4′ C5′ O2′ together with the immediate addenda (OH)3′, (OH)4′ and (CH_2)6′ were placed between two parallel planes and flattened as much as possible. In this respect the furanose and pyranose rings are similar to one another. In the furanose ring, however, this process of flattening of the residue has proceeded even further to involve some degree of distortion from the tetrahedral bond angles The angles which the bonds of (OH)3′ and (OH)4′ make with the adjacent ring bonds are definitely high at 115, 118, 116 and 113°, and this is in a direction such as to bring (OH)3′ and (OH)4′ more nearly into the mean plane of the residue The average bond angle within the ring is 104°. The distances within the furanose ring are as follows.

C2′–C3′ = 1 43_5 A	C5′–O2′ = 1·42 A
C3′–C4′ = 1·45	O2′–C2′ = 1·43
C4′–C5′ = 1·44	

Since the atomic centres were derived entirely from the Fourier projections, these distances are in remarkably close agreement. The average of them (1·43_5) is very definitely lower than a single bond separation, and this fits in well with the observed distortion of bond angles. Evidently the furanose ring is a more closely linked unit than is suggested by the single bonds of the chemical formula The average C-C distance outside the rings is 1·51 A, and the average C-OH distance is 1·41 A The mean deviation of individual bond lengths of one type from the average of that type

is 0·04 A, and this quantity may perhaps also be taken as the probable error to which the bond lengths are subject. The maximum possible error is thought to be about 0·1 A To establish the bond lengths and angles with greater accuracy would require three-dimensional methods, since some of the atoms (e g O 6, O 2′) are not resolved from other atoms in any of the three projections used in this work

An interesting feature of the structure is that each sodium is in contact with one bromine, with a distance between them of 2 94 A This agrees well with the sum of the radii (0 98 + 1·95 = 2·93 A) and is a little smaller than the Na-Br separation in NaBr (2 98 A) In addition to the bond between them the Na and Br atoms in the structure have, of course, other linkages The sodium has bonds to two water molecules and three hydroxyl groups in addition to its bond to Br, the bond lengths being

Na—(OH)6 = 2 42 A	Na—(H_2O)1 = 2 56 A
Na—(OH)6′ = 2·32	Na—(H_2O)2 = 2 41
Na—(OH)4 = 2·62	

The co-ordination group around Na shows a small distortion from a regular octahedral shape The bromine ion has bonds to one water molecule and four hydroxyls in addition to its sodium bond, with the following bond-lengths.

Br—(OH)6 = 3 35 A	Br—(OH)3′ = 3·43 A
Br—(OH)4′ = 3·26	Br—(H_2O)1 = 3 57
Br—(OH)6′ = 3 29	

These agree very well with the accepted radii The group around Br is not of any regular shape

Each of the water molecules possesses four contacts short enough to be regarded as bonds, and these may be taken as usual to consist of two incoming and two outgoing bonds. The disposition of these bonds is not, however, regularly tetrahedral, and in fact in the case of (H_2O)1 is almost flat. The distances involved in addition to those already quoted are

(H_2O)1—(OH)1′ = 2·73 A	(H_2O)2—(OH)3 = 2 99 A
(H_2O)1—(OH)3 = 3 00	(H_2O)2—(OH)2 = 2 67
	(H_2O)2—(OH)3′ = 2·70

so that they all lie between 2 65 and 3 00 A

Besides the bonds discussed above there are two direct contacts between hydroxyl groups of different molecules They are

($OH)_2$—(OH)1′ = 2·77 A and (OH)4—(OH)4′ = 2 66 A

These bonds altogether lead to the very simple rule that every hydroxyl group in the sugar molecule makes *two external bonds*, one outgoing and one incoming Figure 5 is a schematic diagram showing the eight OH groups removed from the molecule, and how they are linked externally to sodium and bromine ions and the water molecules This diagram is, of course, only correct as far as first neighbours are concerned. For instance the bond shown as passing from sodium via (OH)6′

to bromine in reality passes from sodium via the hydroxyl to an adjacent bromine to which the sodium is not directly linked Bond strengths of 1/6 units can conveniently be given to each bond in figure 5

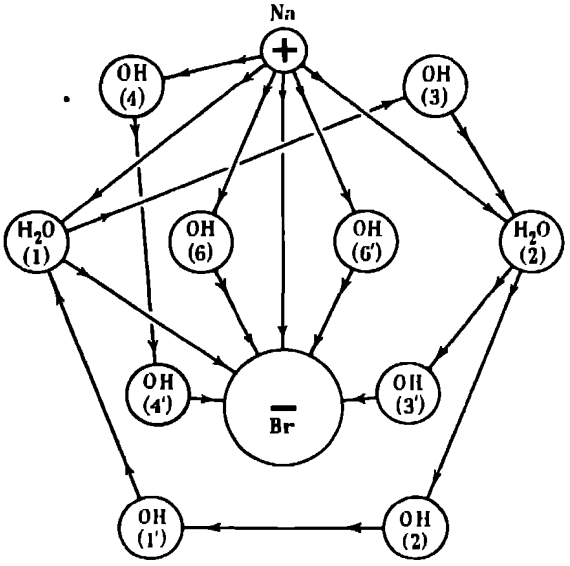


FIGURE 5 Diagram of bond structure

The bond diagram of figure 5 can be used to draw attention to one particular point There is a 'ring' of bonds in the figure, namely

$$(H_2O)\,1—(OH)\,3—(H_2O)\,2—(OH)\,2—(OH)\,1'—(H_2O)\,1$$

which only involves water and hydroxyl dipoles, and which could be reversed in direction so that the bonding would appear to be equally satisfactory In the complete structure the 'ring' of bonds is a spiral of infinite extent in a direction parallel to the c-axis A similar reversible bond chain in Rochelle salt has been suggested by Beevers & Hughes as the explanation of the abnormal dielectric properties of that substance It would be of interest to know whether these bond chains in sucrose sodium bromide are similarly reversible by the action of electric fields If so, this compound would be another example of a ferroelectric, and when polarized would show a reduction of symmetry from orthorhombic to monoclinic

Discussion

The question arises as to whether the structure described in this paper is the only one consistent with the X-ray data, for it has often been pointed out that in structures containing both heavy and light atoms the latter may be moved appreciably without altering radically the calculated F's, so that discrepancies in a table of calculated and observed F's thought to be due to experimental error might actually be due to completely wrong placing of one or more light atoms It is true that in this case the postulated structure fulfils certain stereochemical conditions and gives

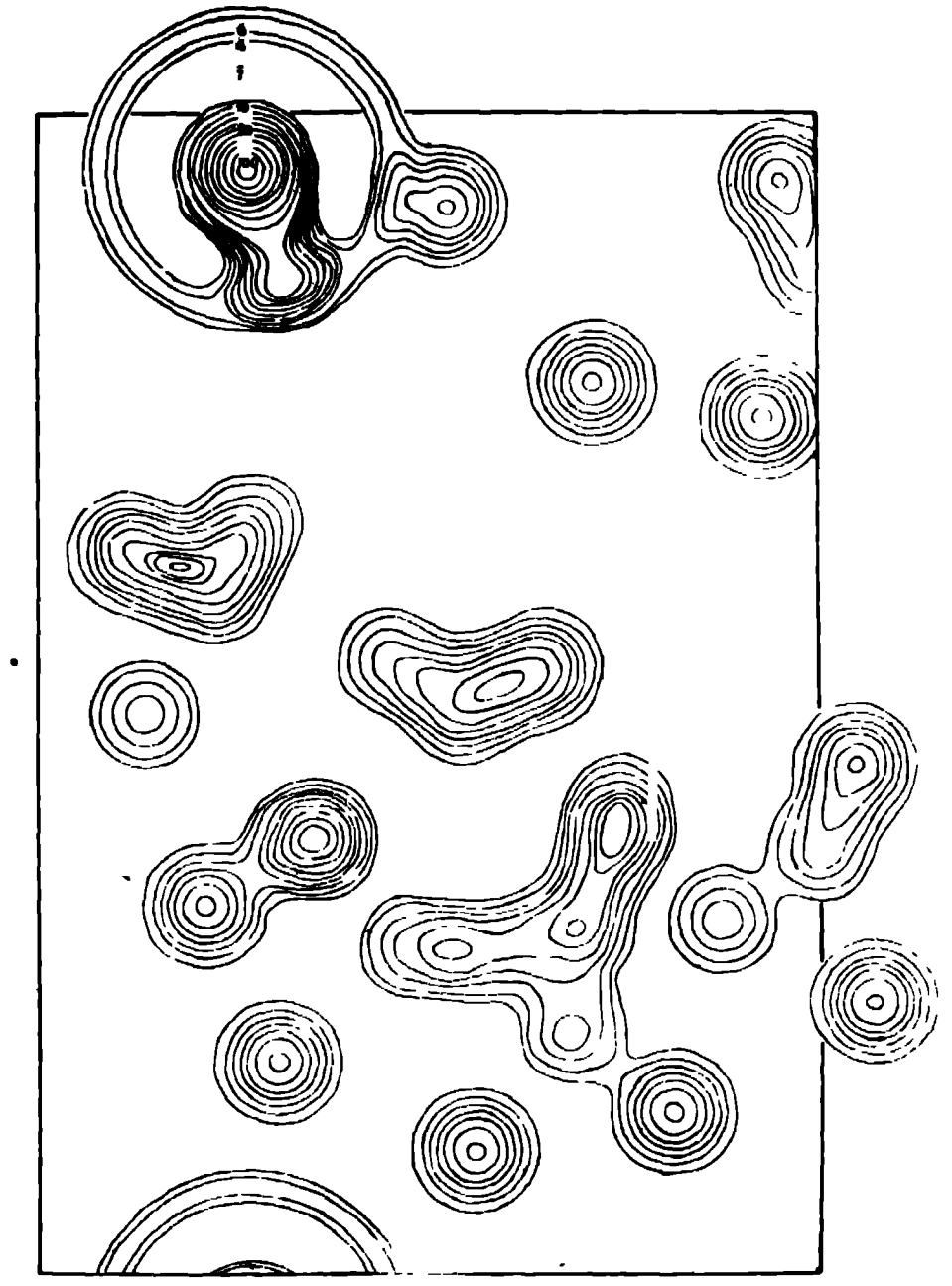


FIGURE 6*a* (Transparent overlay to figure 6.) Electron density projected on to (010) calculated from the structure First contour at $4e/A^2$, then at an interval of $2e/A^2$.

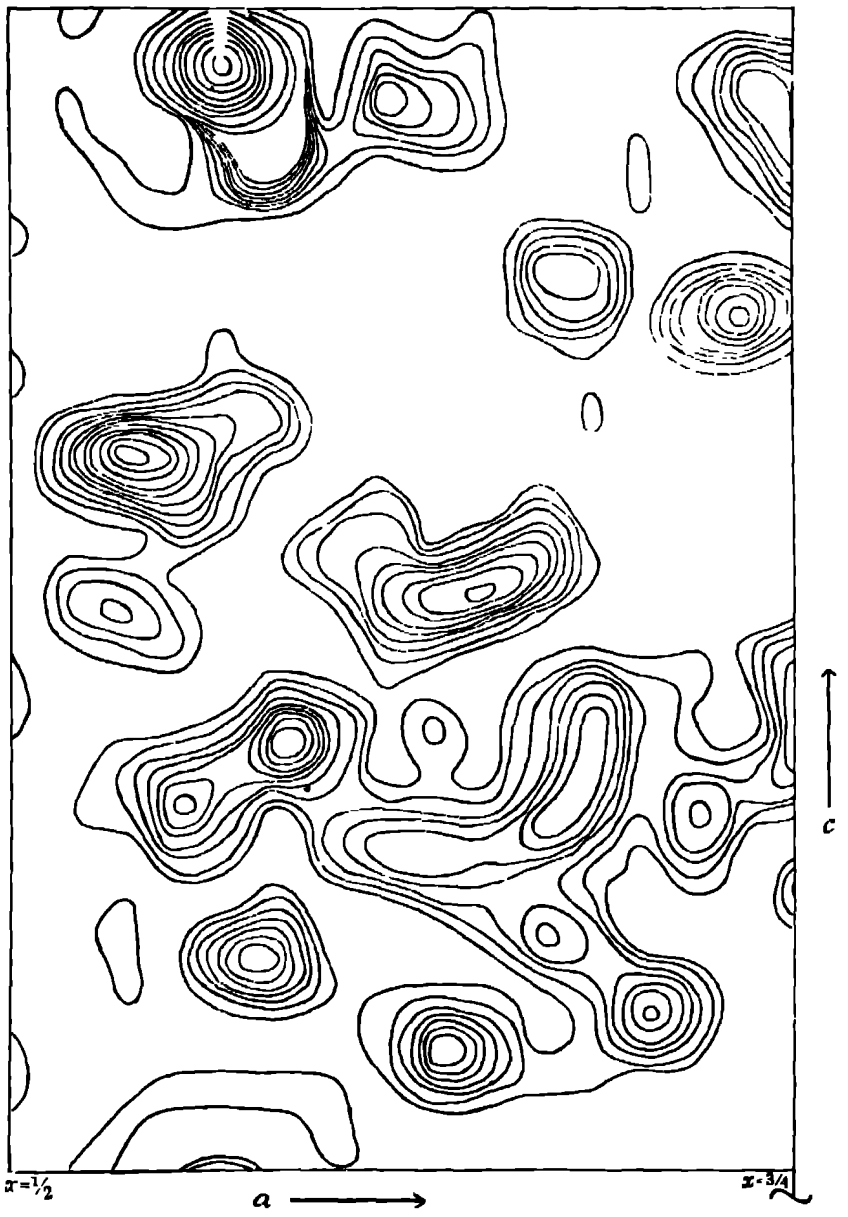


FIGURE 6 Projection of portion of cell from $x = \frac{1}{2}$ to $x = \frac{3}{4}$ on to (010) showing electron density from X-ray reflexions.

extremely good interatomic distances, but it may be enquired how far the final result is independent of all such considerations About 600 zero layer-line structure amplitudes of $C_{12}H_{22}O_{11}$ NaBr $2H_2O$ have been calculated and the agreement with observed values is satisfactory, but it is not proposed to reproduce tables of these here because it is felt that they can be presented in a more compact form which at the same time offers more convincing proof of the correctness of the structure This form is the presentation of calculated and observed electron density distributions—the Fourier transforms of the data contained in the customary tables. This method has the further advantage that the theoretical electron density of a given structure may be derived by a direct method, the amount of computation involved being very much less than in calculating structure factors. That this should be so seems obvious in view of the fact that the electron density in isolated atoms is known and is approximately unchanged by chemical combination, but it must be borne in mind that there would be little point in comparing the observed distribution with that of the Hartree model for example, since observations made with radiation of wave-length 1 54 A do not give the true distribution of electrons within the atom. The distribution given is that of an imaginary atom which has the same scattering factor as the real atom up to a certain limit ($\sin\theta = 1$), after which its scattering factor is zero This is a more diffuse distribution than actually obtains, and particularly in the case of heavy atoms it is not one that falls regularly from a central maximum to zero, but one which after a certain distance from the centre alternates between positive and negative values This may be regarded as a diffraction effect

The apparent electron distribution in different atoms was found by calculating structure factors for a single atom contained in a unit cell of the same dimensions and space group as those of $C_{12}H_{22}O_{11}$ NaBr $2H_2O$ (a simple matter if the atomic co-ordinates are suitably chosen), and then combining those F's which would have been observable with Cu $K\alpha$ radiation in a two-dimensional synthesis to give the apparent electron distribution in that atom Values of the atomic scattering factor were taken from the usual sources, and no account was taken of the presence of the hydrogen atoms The following shows the maximum electron densities values Br (e/A^2), 100·0, Na (e/A^2), 26·8, O (e/A^2), 16 1, C (e/A^2), 10·9 The following table shows the observed maximum electron densities in the projections of the structure of $C_{12}H_{22}O_{11}$ NaBr $2H_2O$ for those atoms which are sufficiently resolved in a projection for this figure to be obtainable

atom	(O)1	(OH)2	(OH)3	(OH)4	(O)5	(OH)1′	(OH)3′
e/A^2 on (001)	15 7	16 5	16 9	—	18 0	16·1	—
e/A^2 on (010)	-	18 5	16 5	17 3	–	16 5	17 3
atom	(OH)4′	(OH)6′	(H_2O)1	(H_2O)2	(CH_2)1′	(CH_2)6′	(CH)1
e/A^2 on (001)	16 5	16 5		13 9	12 7	12 3	12 7
e/A^2 on (010)	16 0	—	14 7	13 0	—	—	—
atom	(CH)2	(CH)3	(CH)4	(CH)5	(CH_2)6	Br	
e/A^2 on (001)	11·6	—	12 0	10 1	9 4	90 0	
e/A^2 on (010)	13·0	13 0	—	—	10·4	95·0	

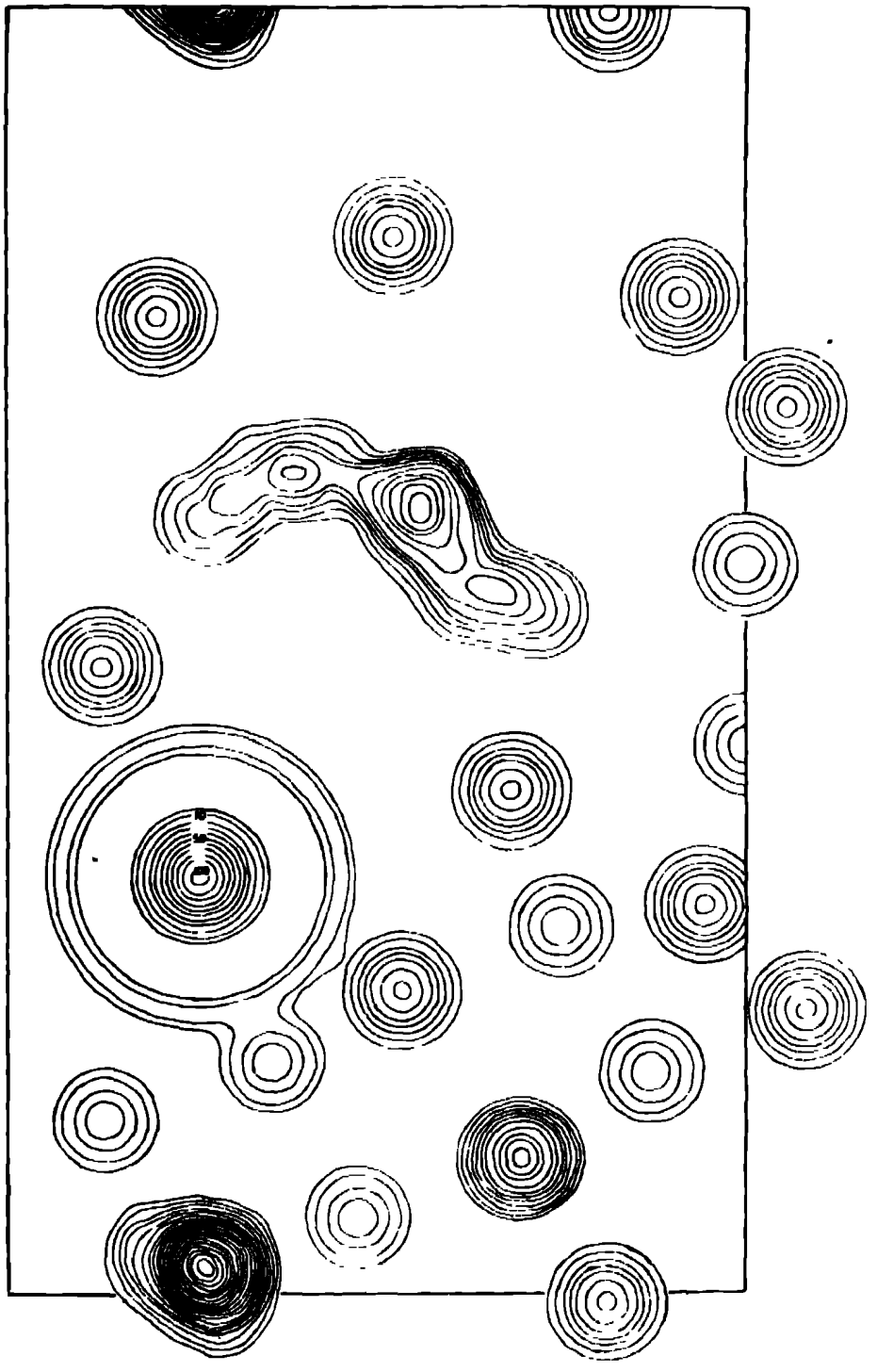


FIGURE 7a. (Transparent overlay to figure 7) Electron density projected on to (001) calculated from the structure

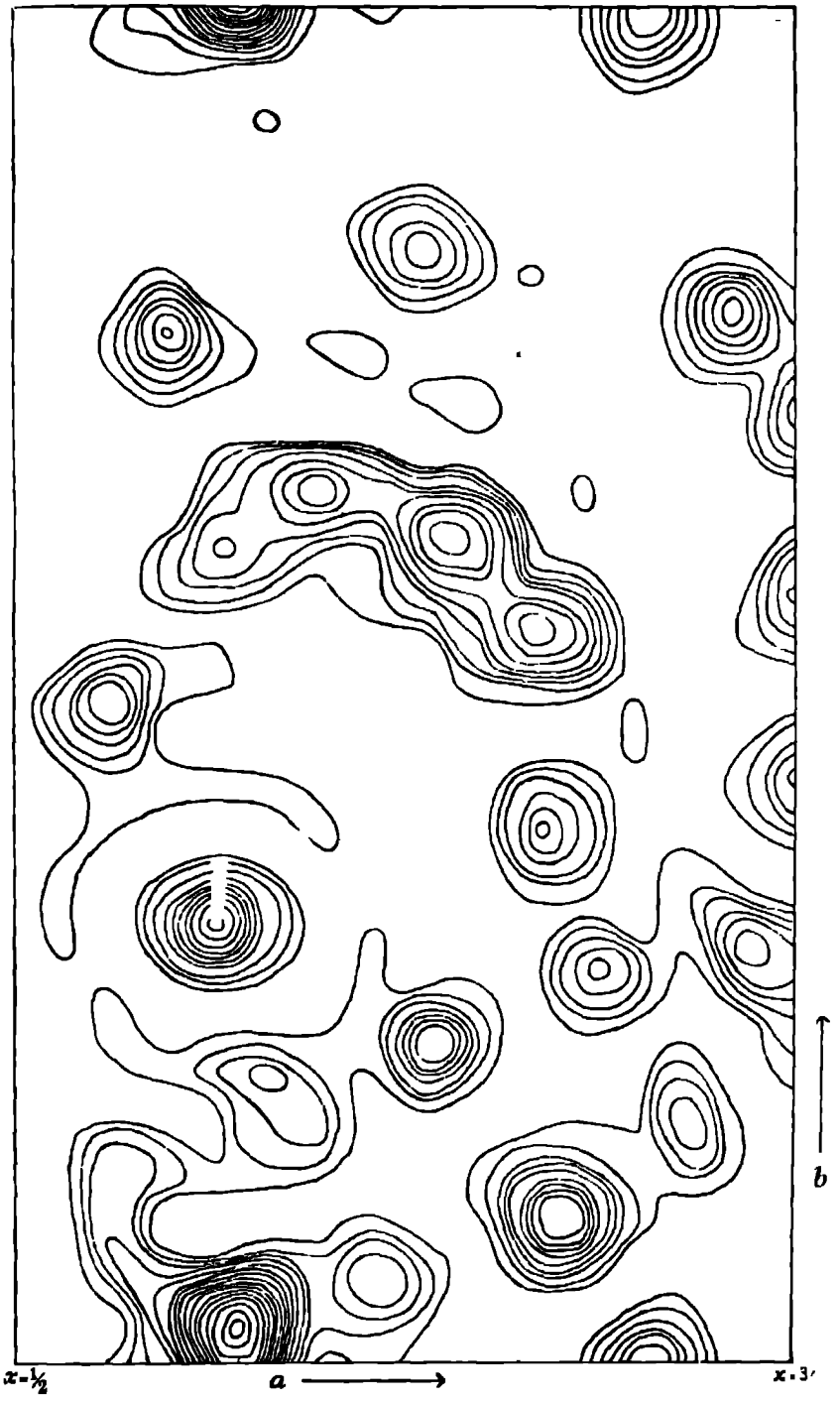


FIGURE 7. Projection on to (001) showing electron density from X-ray reflexions

The average maximum electron densities are 16 8 for O or OH, 13·9 for water molecules, and 11·7 e/A^2 for carbon atoms, CH or CH_2 groups, the mean deviation from the average being 0·8 e/A^2 and the maximum deviation 2 3 e/A^2. The average value of the observed 'background' electron density is + 1·3 e/A^2, and the mean deviation from the average 1·0 e/A^2, points in the neighbourhood of the bromine atom being excluded because of the marked diffraction effects in this region These results indicate that the observed electron density is subject to a probable error of about 1·0 e/A^2 and the maximum error must be about 4 e/A^2, since small ghost peaks of this height appear The observed electron density of the water molecules is thus consistently lower than that calculated for an oxygen atom by an amount which is too great to be due entirely to experimental error. Possibly this is a real effect due to the water molecules having a greater thermal movement than atoms which form part of the molecule In figures 6 and 7 the calculated electron density at all points of the unique areas of the (010) and (001) projections is compared with the observed values It will be seen that the electron distribution is very similar in both and that no atom could be moved appreciably without destroying the agreement It may therefore be claimed that the structure is based entirely on the X-ray data

References

Cox, E G , Goodwin, T. H & Wagstaff, A I 1935 *J Chem Soc* p 1495
Cox, E G & Jeffrey, G A 1939 *Nature*, **143**, 894
Cochran, W 1946 *Nature*, **157**, 231

Atmospheric tides in the ionosphere
II. Lunar tidal variations in the F region near the magnetic equator

By D. F. Martyn, D.Sc

Australian Council for Scientific and Industrial Research

(*Communicated by Sir Edward Appleton, F.R.S.—Received 6 September* 1946)

Three years' data for the heights and critical frequencies of the F_1 and F_2 regions at Huancayo, Peru, are examined for lunar tides. Semi diurnal lunar tides are found in all quantities save $f^0_{F_1}$. It is found that the lunar variation in the F_2 region depends on solar time in both phase and amplitude. At certain solar epochs it attains very large amplitudes, up to 60 km in $h^{max}_{F_2}$ and 28 % in $f^0_{F_2}$.

The theory of these variations, originated in a previous paper, is further developed.

1. Introduction

In a previous paper (Martyn 1947), afterwards referred to as I, a theory was propounded that the major anomalies in behaviour of the F_2 region of the ionosphere were due to the action of the solar atmospheric tide, which tidal theory indicates may have a large amplitude at this level. In brief, the theory showed that the tidal motion, though predominantly horizontal, would cause large vertical movements of ionization, because of the presence of the earth's magnetic field. The existence of gradients in this vertical velocity would distort the shape of the ionized region, so that both the maximum electron density N(max.), and h(max.), the height at which this occurred, would depart notably from the values to be expected for a 'Chapman region'.

In I certain data were presented showing the existence of large 12-hourly harmonics in h and N, and it was further shown that the theory could qualitatively explain some of the major anomalies in F_2. This evidence by itself, however, is hardly sufficient to establish the validity of the theory. The isolation of solar *tidal* effects in the F region is rendered difficult by the fact that the sun also exerts a strong *ionizing* effect, with a 24 hr. periodicity, but with appreciable 12-hourly harmonics. These may be expected to give rise to 12-hourly harmonics in N(max.) and h(max.). Appreciable effects of similar periodicity may also arise from thermal heating of the F region by the sun. For these reasons it was considered desirable in I to study, particularly, data for the hours of darkness, and to show that there occurred then distinct temporal maxima in N(max.) and h(max.), displaced 12 hr. from the daytime maxima.

The study of the effect of tidal motions in the F region would obviously be much facilitated if it were possible to observe a lunar tide at this level, since the effect of this tide would presumably be uncomplicated by ionizing and thermal influences.

of lunar periodicity Now if the present theory of F_2 anomalies is valid then the solar tide causes changes in N(max) of the order of 50 % in certain cases, and changes in h(max.) of the order of 50 to 100 km The lunar tide at the ground is between a tenth and a fifteenth of the solar, so that, if a like ratio is preserved in the upper air, quite observable lunar effects in the F region may be anticipated

With these considerations in mind the F region data from a number of ionospheric observatories have been examined for lunar periodicities In the case of the F_2 region these have been found in both N(max) and h(max), at all sites so far examined The largest effects so far investigated are found at Huancayo, Peru, near the magnetic equator. Although the results for this site show some other features not found at higher latitudes, it has been thought advisable to present them first, on account of their magnitude, leaving those for other sites to later papers

2 Data and treatment

The data used were the hourly values of $f^0_{F_2}$, $h^{\max}_{F_2}$, h'_{F_2}, $f^0_{F_1}$ and $h^{\max}_{F_1}$ at Huancayo (lat 12·0° S, long 75 3° W) for the years 1942–5 However, $h^{\max}$ data were not available after June 1944, when the practice of scaling this quantity was generally abandoned, except in Australia The values of f^0, the critical penetration frequency of a region, are proportional to $\sqrt{N}$(max), and are accurate to about 0·1 Mcyc /sec The readings of h are accurate to about 10 km.

All data were made available by the Director of the Department of Terrestrial Magnetism and Electricity of the Carnegie Institution of Washington, which maintains the Huancayo Observatory The methods used in deriving the lunar variations follow those developed by Chapman & Bartels (1940) for the determination of the lunar barometric tide, and the lunar magnetic variations. Following these authors, the diurnal solar variation was removed from the data before they were rearranged in lunar time, reckoned from lower transit For convenience, times of lower transit at Greenwich were used, as given in the *Nautical Almanac* A small correction, amounting to $nL^0/29\ 5$, where L^0 is the longitude of Huancayo measured westwards, and n is the order of the harmonic under consideration, was later applied to the phase of the harmonic, in order to bring the results to true local lunar time The data for each calendar month was arranged in lunar time on one sheet and added up. Groups of months were then subjected to harmonic analysis, as described later The results of these analyzes were plotted on harmonic dials The groupings of points on the dials show the degree of consistency of the results, and can be used to estimate the probable error if a clear tidal effect be found.

3. The F_1 region

(3·1) *Lunar tidal variations in* $h^{\max}_{F_1}$

In figure 1 are plotted the results for $h^{\max}_{F_1}$ for the 30 months for which data were to hand The point at each lunar hour on this graph was obtained by adding the monthly totals for that hour and dividing the grand total by the number of obser-

vations The latter number was approximately one-third of the number of the lunar days in the 30 months, since the F_1 region is only present for about a third of each day Each point on the graph therefore is derived from 366 observations The graph appears to show a semi-diurnal variation of half-amplitude (P_2) about half a kilometre, the maximum occurring about 04 hr. lunar time (t_2)

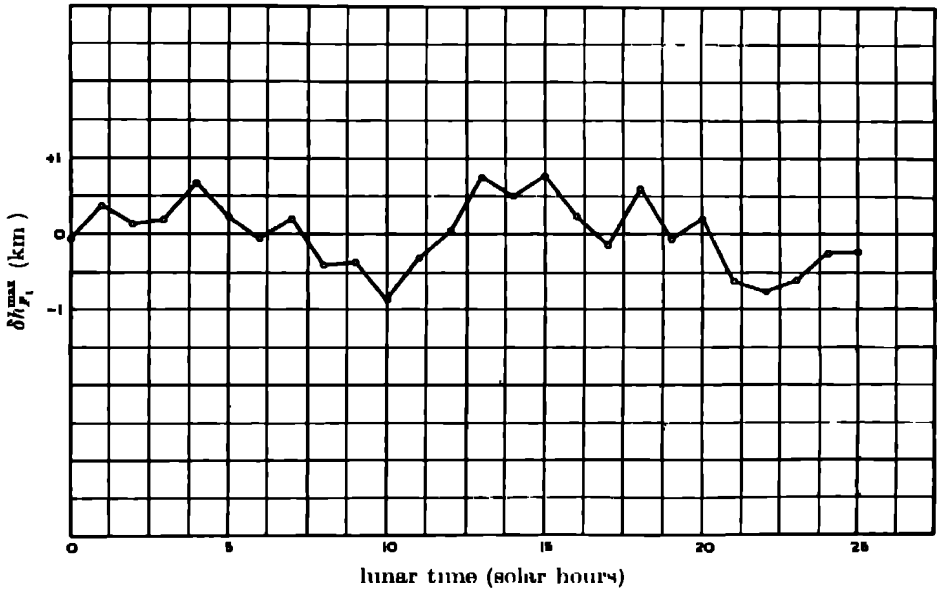


FIGURE 1 Average lunar variation of $h'^{max}_{F_1}$ at Huancayo, 1942-4.

To test this conclusion more rigorously the data were divided into groups of months and the average variation for each group subjected to harmonic analysis The groups selected were (*a*) January, February, November, December, (*b*) May, June, July, August and (*c*) March, April, September, October, i e the two solstitial groups and the equinoctial group This was done for each year separately. Since only 6 months' data were available for 1944 the results for each group in this year carry only half the weight of those for the other years

The diurnal harmonic in these groups is found to be of small amplitude and apparently random phase and is not considered further.

The results for the semi-diurnal harmonic are plotted on the 12 hr harmonic dial in figure 2, the semi-amplitude P_2 being plotted as distance from the centre of the dial, and t_2 as time on the dial regarded as a clock face. Group (*a*) months are plotted as open circles, group (*b*) as filled-in circles and group (*c*) as crosses A short fine line is drawn across the 1944 points, which carry only half the weight of the others The grouping of the points shows the reality of a semi-diurnal variation of half-amplitude approximately 0 5 km , attaining a maximum at about 04 lunar hours. There is some evidence that the variation is larger in summer than in winter, but the

amount of scatter in the points indicates that the data used are too few to permit more detailed conclusions. This is not surprising when it is recalled that the accuracy of each individual measurement of h^{max} is only 10 km and the data cover less than one effective year, since the F_1 region can only be observed in daylight

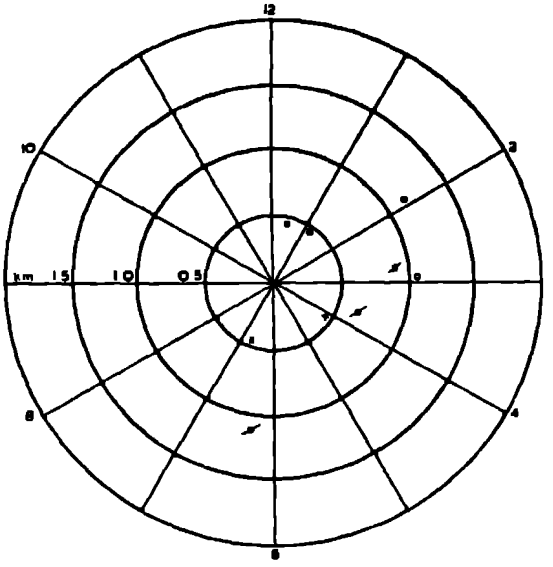


FIGURE 2 Harmonic dial of semi-diurnal lunar variation in $h'^{max}_{F_1}$ at Huancayo, 1942-4. O Jan., Feb., Nov., Dec., × Mar., April, Sept., Oct., ● May, June, July, Aug., ⁄ indicates plot of half-weight

(3·2) *Lunar tidal variation in* $f^0_{F_1}$

The data for the full three years 1942-4 in $f^0_{F_1}$ have been subjected to analysis similar to that outlined in (3·1) No evidence of a consistent lunar variation has been found If such exists it must have a half-amplitude considerably less than 0·01 Mcyc /sec

4 THE F_2 REGION

(4·1) *Lunar tidal variations in* h'_{F_2} *and* $h'^{max}_{F_2}$

In figure 3 are plotted the average lunar tidal variations of h'_{F_2} (dotted line) and $h'^{max}_{F_2}$ (full line). Each point on the h'_{F_2} curve is derived from 1039 hourly readings and each of those on the $h'^{max}_{F_2}$ curve from 872 hourly readings

There is clearly a large semi-diurnal variation of several kilometres The coefficients of the first four harmonics are shown for both quantities in table 1, where P_1, P_2, etc., represent the half-amplitude of the 1st, 2nd, etc., harmonic in km, and t_1, t_2, etc., the lunar time of maximum amplitude, reckoning 24 lunar hours to the lunar day.

There is therefore a large semi-diurnal variation, which is more than twice as large in h'^{max} as in h'. The phase of the variation in the latter is retarded almost an

hour on that in the former. The only other harmonic of comparable importance is P_1.

These two harmonics were further investigated by means of the harmonic dial. The procedure adopted was identical with that described in (3 1) for $h^{\max}_{F_1}$

TABLE 1 HARMONIC COEFFICIENTS OF LUNAR VARIATION IN h'_{F_1} AND $h^{\max}_{F_1}$ AT HUANCAYO

	P_1	P_2	P_3	P_4	t_1	t_2	t_3	t_4
h'_{F_1}	0·51	2 32	0 058	0 48	22 2	9 23	6 26	3 90
$h^{\max}_{F_1}$	1 46	5 20	0·04	0 26	22 6	8 43	5 56	2 90

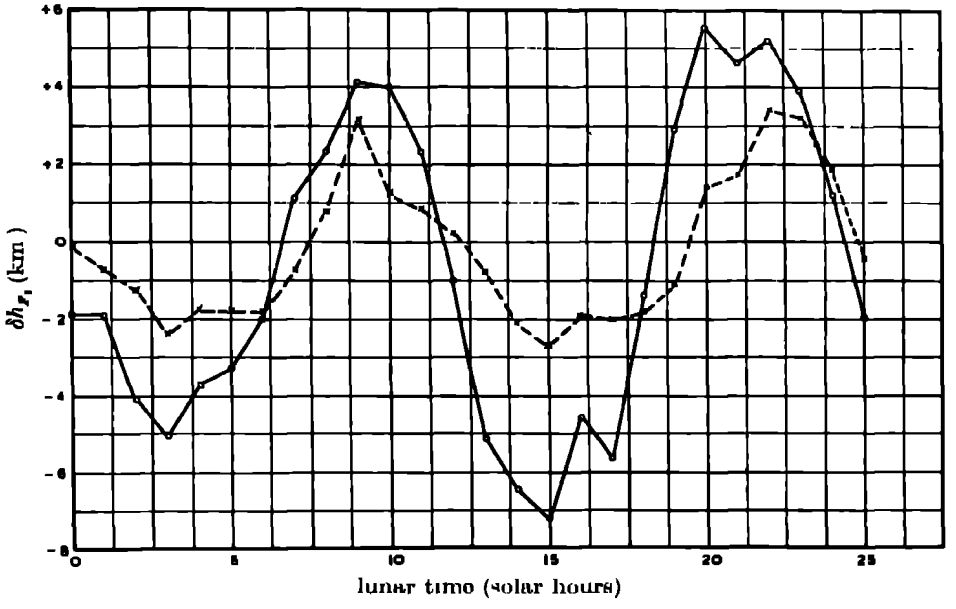


FIGURE 3 Average lunar variations of $h^{\max}_{F_1}$ (full line) and h'_{F_1} (broken line) at Huancayo, 1942–4

The results for P_1 are shown in the diurnal dial in figure 4 The h' points are indicated by filled-in circles and the $h^{\max}$ points by crosses Each point represents a group of 4 months as in (3 1). The scatter of the points is such that it is concluded that there is no evidence for the existence of a diurnal lunar variation in h

The results for P_2 in $h^{\max}_{F_1}$ and h'_{F_1} are plotted in figures 5 and 6 respectively In these figures, since a real variation is obviously present, one reverts to the plan described in (3 1) of representing each group of months (*a*) to (*c*) by a distinctive plot (open circles, filled circles and crosses) In both figures no points are found to fall in the half-circle representing 0 to 6 hr , so space is saved by omitting this half of the dial

Clearly there is a real semi-diurnal variation in both quantities, that in h^{max} being rather more than double that in h' The phase of the variation in h^{max} is about 1 hr in advance of that in h' It is also clear that the variation in group (a) months is about twice that in group (c) months for h', and for h^{max}, rather more than four

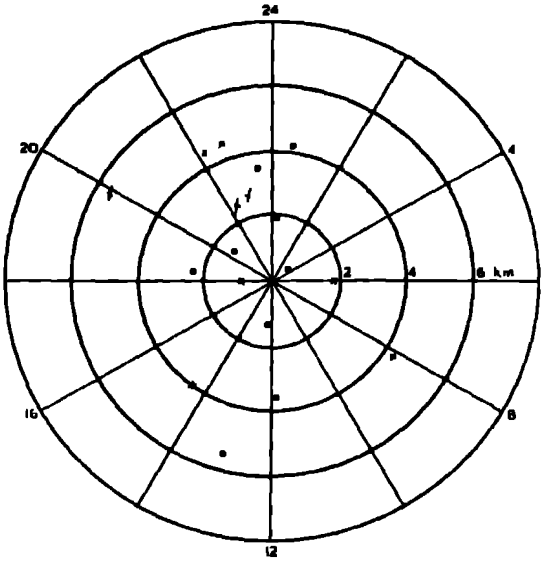


FIGURE 4. Harmonic dial of diurnal variation $h_{F_1}^{max}$ ×, and h'_{F_1} ●, at Huancayo, 1942–4. Points thus ⁄ carry only half-weight

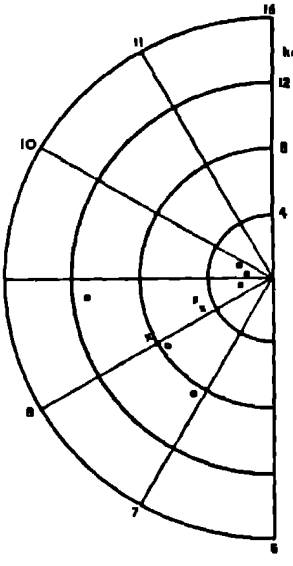


FIGURE 5

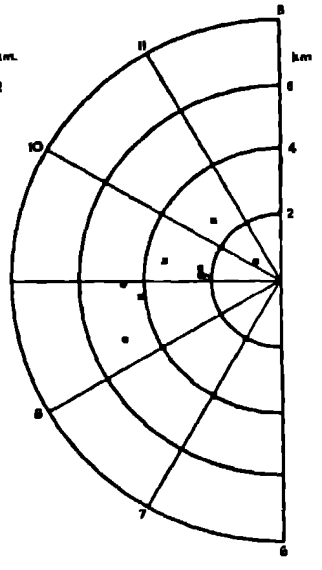


FIGURE 6

FIGURES 5 and 6 Harmonic dials for semi-diurnal lunar variations of $h_{F_1}^{max.}$ (figure 5) and h'_{F_1} (figure 6) at Huancayo, 1942–4. Note that scale of figure 6 is double that of figure 5.

times as great The equinoctial amplitudes occupy an intermediate position between those for group (*a*) and group (*b*) months The more precise determination of these seasonal ratios must await the analysis of considerably more data

(4·2) *Lunar tidal variations in* $f^0_{F_1}$

The average lunar variation in $f^0_{F_1}$ for the full 3 years 1942–4 is shown in figure 7. Harmonic analysis of this variation gives the results shown in table 2

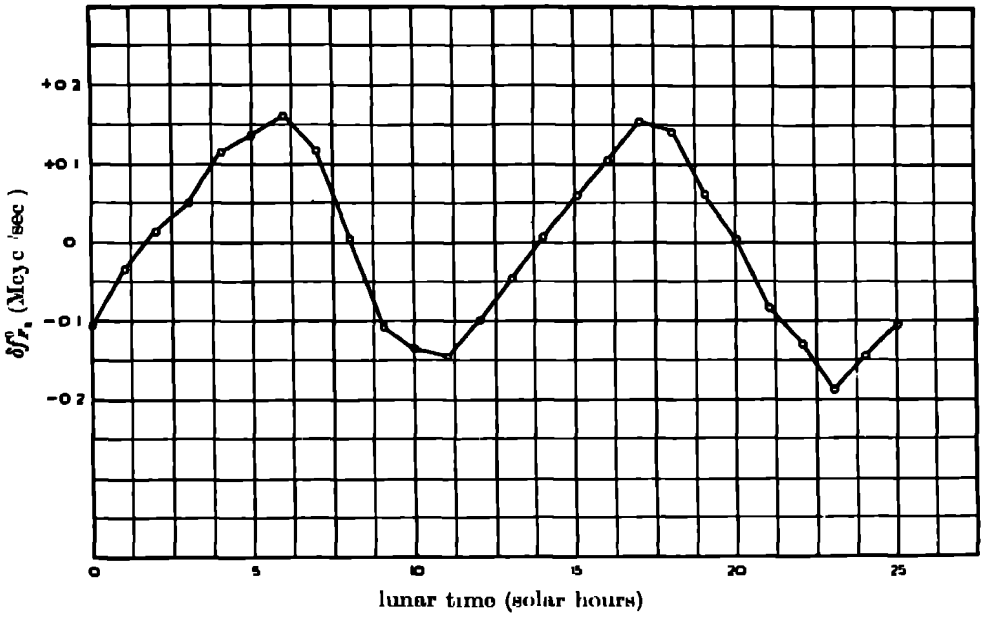


FIGURE 7 Average lunar variation of $f^0_{F_1}$ at Huancayo, 1942–4.

TABLE 2 HARMONIC COEFFICIENTS OF LUNAR VARIATION IN $f^0_{F_1}$ AT HUANCAYO
P'S IN MCYC /SEC , *t*'S IN LUNAR HOURS

P_1	P_2	P_3	P_4	t_1	t_2	t_3	t_4
0 018	0 10	0 013	0 02	9 78	4 33	6 50	5 94

There is clearly a strong semi-diurnal harmonic of half-amplitude 0·10 Mcyc./sec.

This conclusion is confirmed by the use of the harmonic dial in a manner similar to that described in (4 1) In this case all the points are found to fall in the quadrant between 3 and 6 hr , so for economy the remainder of the dial is not shown in figure 8 It appears likely that the variations in f^0 should have greater physical significance if expressed in fractional form Thus, since $N(\text{max.}) \propto f^{0^2}$, it is seen that

$$\delta N/N = 2\delta f^0/f^0$$

For this reason the quantity plotted in figure 8 is $200P_2/\text{mean}f^0_{F_1}$, which is the half-amplitude of the variation in N(max), expressed as a percentage The results of the analysis of 1945 data became available as this paper was nearing completion, and

are included in figure 8, the points in question being labelled with a 5. The amplitudes of variations in the monthly groups (*a*) and (*c*) seem to be distinctly larger than those in group (*b*), and there is evidence that the phase of the (*a*) months variation is about an hour in advance of that for the equinoctial months.

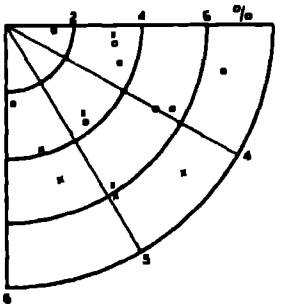


FIGURE 8 Harmonic dial of semi-diurnal lunar variation of $f^0_{F_1}$ at Huancayo, 1942 5
Quantity plotted is $200P_2/\text{mean } f^0_{F_1} = 100 \times \frac{1}{2}\delta N_m/N_m$

(4·3) *Influence of solar time on F_2 lunar variations*

Comparison of the results in (3 1) with those in (4 1) shows that there is a large difference in phase between the semi-diurnal lunar variations at the heights of the F_1 and F_2 regions Moreover, comparison of figures 5 and 6 shows that there are differences in phase and amplitude of the variations at the two levels $h^{\max}$ and h' in the F_2 region itself This suggests that there may be notable changes with height in both the phase and amplitude of the lunar atmospheric oscillation Now the F_2 region at Huancayo undergoes very large variations of height, of more than 100 km , in the course of a solar day It seems desirable therefore to examine whether the F_2 lunar variations show any dependence on solar time For this purpose the data for $h^{\max}_{F_2}$ for each solar hour separately was analyzed for the lunar semi-diurnal variation In the case of f'_{F_2}, experience indicated that adequate information might be obtained by analyzing only each odd solar hour, 1, 3, 5, , 23 Later on it was found desirable to supplement this by analysis of the even hours 4, 6 and 24, which proved to be times of rapid variation.

The results for $h^{\max}_{F_2}$ are exhibited in figure 9 which shows P_2 (curve I) and t_2 (curve II) for each solar hour Figure 10 shows the same quantities in the analysis of $f^0_{F_2}$. In this case, for the reason mentioned in (4 2), the amplitude P_2 at each solar hour was divided by the average value of $f^0_{F_2}$ at that hour, and expressed as a percentage It is seen at once from these figures that the phase and amplitude of the semi-diurnal lunar variations in F_2 at Huancayo are markedly dependent on solar time The half-amplitude of the variation in $h^{\max}_{F_2}$ ranges from 0·34 km at 02 hr to as much as 16 km at 16 hr., while the half-amplitude density variations in N(max.) range from 14 6 % at 05 hr. to 2 % at 07 hr In their strong dependence on solar time the lunar F_2 variations at Huancayo are markedly different from those at other sites at present being analyzed

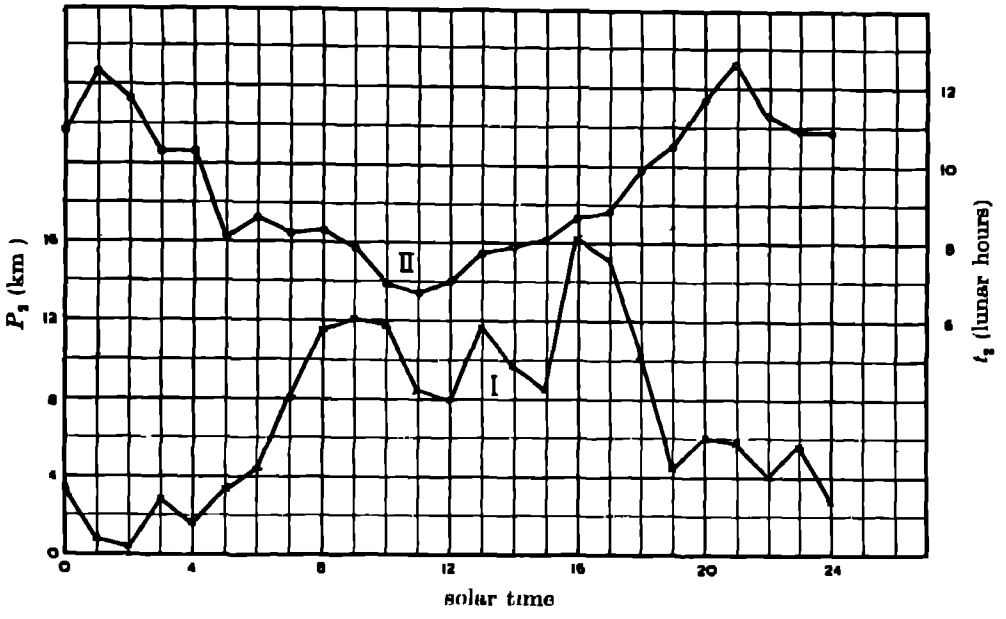


FIGURE 9 Variation with solar time of the lunar harmonic coefficients P_2 (curve I) and t_2 (curve II) for $h'^{max}_{F_2}$, Huancayo.

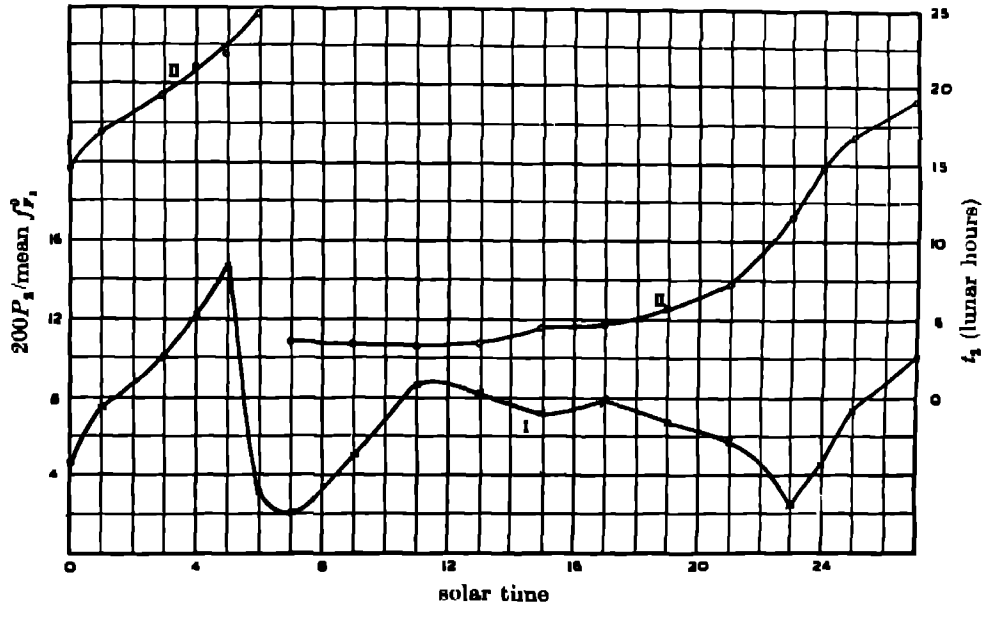


FIGURE 10 Variation with solar time of the lunar harmonic coefficients $200P_2$/mean $f^0_{F_2}$ (curve I) and t_2 (curve II) for $f^0_{F_2}$, Huancayo.

To illustrate the remarkable amplitude of the lunar variation in $h_{F_1}^{max}$ at certain solar epochs, the variation for solar hour 17 in the group (a) months has been selected. Figure 11 shows the average lunar variation at these times, and is therefore derived from only 10 lunar days Harmonic analysis shows the semi-diurnal component to have an amplitude of 56 km

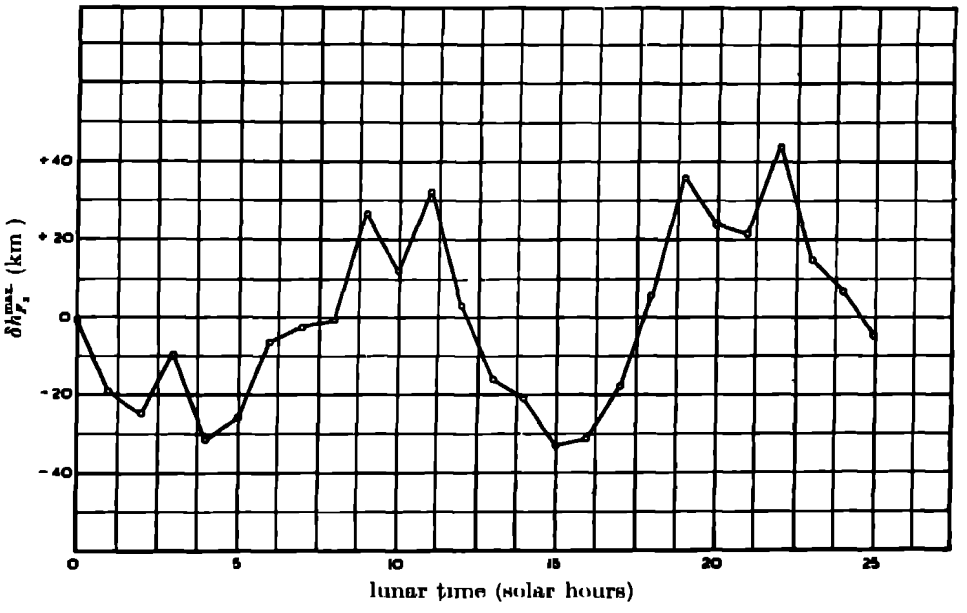


FIGURE 11. Average lunar variation of $h_{F_1}^{max}$ at 17 hr. solar time for months Jan , Feb , Nov , Dec. 1942–4, Huancayo

5 DISCUSSION OF RESULTS

(5 1) *General significance of results*

The lunar variation found in $h_{F_1}^{max}$ is of the same order of magnitude as that found for h'_E by Appleton & Weekes (1939) in England The phase of the variation is almost in opposition to that found for the E region however

The most remarkable features of the F_2 variations are their amplitude, and the peculiar dependence on solar time. It is certainly astonishing that the moon can produce variations up to as much as 56 km in the height of this region.

It would be extremely difficult to interpret these results as a direct consequence of the height and pressure variations associated with the lunar atmospheric tide Even if it is assumed that the amplitude and phase of the latter are functions of height it would still seem to be impracticable to explain the complexities the results reveal For example, figure 10 shows that large changes in the amplitude of the lunar variations in N(max) occur between 0 and 5 hr solar time, although the height of the region changes very little between these hours (figure 12)

The magnitude of the lunar variations revealed above lends strong support to the theory outlined in I that the anomalous diurnal, seasonal and geographical, variations in F_2 are caused by the solar atmospheric tide, which is more than ten times larger than the lunar at the ground It seems hardly likely that the solar tide should chance to be less than the lunar over the whole range of heights involved in the F region Moreover, it will be found below that the 'lunar-solar' variations outlined in (4 3) are readily explained on the assumption that there is present in this region a solar tide of considerably greater amplitude than the lunar.

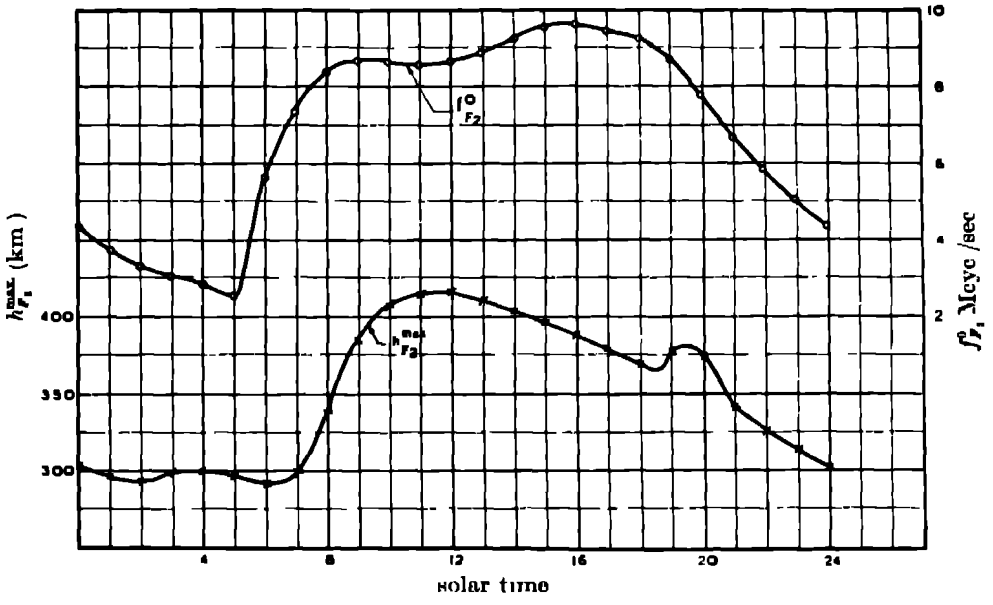


FIGURE 12. Average solar diurnal variations of $f^0_{F_2}$ and $h'^{max}_{F_2}$, Huancayo, 1942–4.

In I the theoretical consequences of the variation of tidal amplitude with height were specially considered Such variation leads to distortion of an ionized region, of a kind such that the variations of $h^{max.}$ and $N^{max.}$ are mutually in phase, or out of phase exactly Now the variations in these two quantities found at a number of locations are normally more nearly in phase quadrature This can be readily seen in the present results by comparing figures 5 and 8.

Now comparison of the results for $h'^{max.}_{F_1}$, h'_{F_1} and $h'^{max}_{F_2}$ suggests strongly that there is a variation of phase with height in the lunar atmospheric tide This is further suggested by some unpublished calculations by Mr G Odgers, who has examined the theory of Pekeris (1937) in this respect.

The effect on an ionized layer of a vertical motion of ions with a velocity varying harmonically with time, and with a phase changing linearly with height is examined in the following section.

(5·2) *Effect on an ionized region of vertical transport of ions with velocity varying harmonically with time and changing phase linearly with height*

Let $v = v_0 \sin(\omega t + \sigma z)$ be the vertical velocity of an ion, measured positively downwards. As in I, ions of both signs drift in the same sense with equal velocities, under the influence of the electrodynamical forces occasioned by the *horizontal* tidal force acting in the presence of the earth's magnetic field. z is the vertical height above an arbitrary datum level, both z and v are measured in units of H_0, the 'height of the homogeneous atmosphere'.

The equation of continuity is then

$$\frac{\partial N}{\partial t} = \frac{\partial (Nv)}{\partial z} \tag{1}$$

The general solution of equation (1) is

$$N = \frac{v_0}{\omega - \sigma v} \psi\left(t + \int \frac{dx}{\sigma v_0 \sin x - \omega}\right), \tag{2}$$

where $x = \omega t + \sigma z$ and ψ is any function. A solution which gives a satisfactory approximation to the form of an ionized region is

$$N = \frac{1}{1 - \kappa \sin x} \cos\left\{\omega t + \frac{2}{\sqrt{(1-\kappa^2)}} \arctan \frac{\kappa - \tan \frac{1}{2}x}{\sqrt{(1-\kappa^2)}}\right\}, \tag{3}$$

where $\kappa = \sigma v_0/\omega < 1$. Figure 13 shows the ionized region represented by equation (3) for the four epochs $\omega t = 0$, $\frac{1}{2}\pi$, π, $\frac{3}{2}\pi$, and for $\kappa = 0{\cdot}3$. It will be observed that the oscillations of $\sigma z_m (\sigma z(\max))$ and $N_m (N(\max))$ appear to be almost in quadrature. When $\kappa \ll 1$ this is exactly the case, and N oscillates between $1+\kappa$ and $1-\kappa$, the maxima and minima occurring at $\omega t = \frac{1}{2}\pi$ and $\frac{3}{2}\pi$ respectively. z_m then oscillates between v_0/ω and $-v_0/\omega$, the maxima and minima occurring at $\omega t = 0$ and π respectively.

The variation of N_m and σz_m with time is shown in figure 14 for $\kappa = 0{\cdot}3$. It will be observed that N_m varies almost sinoidally with a period slightly greater than $2\pi/\omega$. The variation of σz_m departs more markedly from the sinoidal form, the rise in height being rapid and the fall slow. There is also a steady downward trend in height. It will be noted that the maximum in z_m occurs about 130° *ahead* of the N_m maximum. When $\kappa < 0$ ($\sigma < 0$), the phase of the z_m variations lags a similar amount *behind* that of N_m.

In general, $$N_m = (1 - \kappa \sin x_m)^{-1} (1 + \kappa^2 \cos^2 x_m)^{-\frac{1}{2}}, \tag{4}$$

where $x_m = \omega t + \sigma z_m$, so that, when κ is $\ll 1$,

$$N_m = 1 + \kappa \sin x_m \tag{5}$$

Also, when $\kappa \ll 1$, $$z_m = \frac{\kappa}{\sigma} \cos x_m. \tag{6}$$

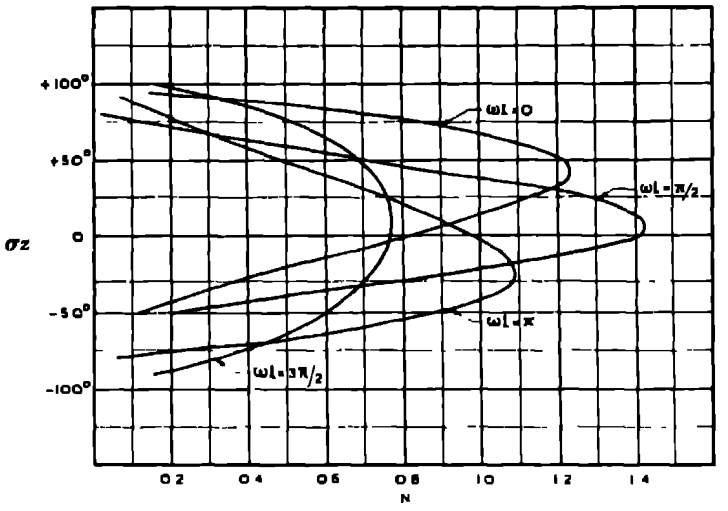


FIGURE 13 Theoretical variations of ionized layer acted upon by harmonically varying vertical velocity whose phase advances linearly with height ($\sigma v/\omega = 0\ 3$).

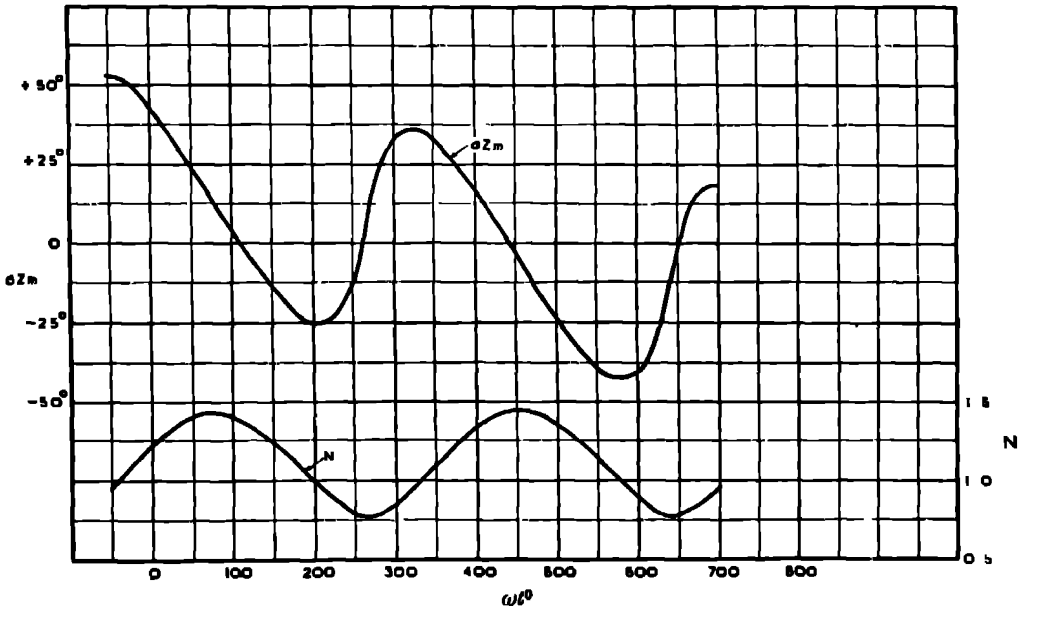


FIGURE 14. Theoretical temporal variations of N_m and σz_m for ionized layers in figure 13

(5·2·1) *Effect on an ionized region of superposition of two velocities of vertical transport*

The suffixes 1, 2 are used to refer to the vertical ionic drifts associated with the solar and lunar tides respectively. Then

$$v = v_1 \sin(x_1 + \theta) + v_2 \sin x_2, \tag{7}$$

where θ is the (unknown) phase difference between the solar and lunar ionic drift velocities Assume that $v_2 \ll v_1$, in accordance with direct measurements at the ground, and the indirect evidence of the ratio of the lunar and solar currents in the ionosphere responsible for the magnetic variations. Then

$$v = v_1\left(1 + \frac{v_2}{v_1}\cos(x_2 - x_1 - \theta)\right)\sin(x_1 + \theta + \phi), \tag{8}$$

where $$\tan\phi = \sin(x_2 - x_1 - \theta)\, v_2/v_1$$

The compound given by equation (8) may be regarded as an harmonically varying velocity, subjected to a small and slow amplitude and phase modulation. Applying equation (5), and neglecting ϕ, then

$$N_m = 1 + \frac{\sigma_1 v_1}{\omega_1}\sin(\omega_1 t + \sigma_1 z_m + \theta) + \frac{\sigma_1 v_2}{\omega_1}\sin(\omega_1 t + \sigma_1 z_m + \theta)\cos(\overline{\omega_2 - \omega_1}.t + \overline{\sigma_2 - \sigma_1}\ z_m - \theta) \tag{9}$$

The first term in equation (9) represents the N_m variation due to the solar tide, the second may be regarded as a lunar variation, the phase and amplitude of which is dependent on the solar tide

The method of treating the data in (4·3) involves the selection of a particular solar hour, or epoch of the solar tide, and the examination of the lunar variations of N_m and z_m around the mean values of those quantities at that epoch.

If the solar variation is written as

$$\nabla N_m(S) = \frac{\sigma_1 v_1}{\omega_1}\sin S, \tag{10}$$

then the lunar-solar variation is

$$\nabla N_m(LS) = \frac{\sigma_1 v_2}{\omega_1}\sin S\cos(\omega_2 t + \sigma_2 z_m - S) \tag{11}$$

Notice that the LS variation retards in phase at the rate the S variation advances, while its amplitude varies as $\sin S$.

By equation (6) $$\nabla z_m(LS) = \frac{v_2}{\omega_1}\cos S\cos(\omega_2 t + \sigma_2 z_m - S). \tag{12}$$

(5·3). *Deductions from the results*

The average solar diurnal variations of $f_{F_1}^0$ and $h_{F_1}^{\max}$ at Huancayo for the years 1942–4 inclusive are shown in figure 12 The general shape of the $z_m(S)$ variation revealed here suggests that $k \nless 1$. The half-amplitude of $\nabla z_m(S)$ appears from these data to be about 50 km., which by equation (6) is v_1/ω_1, so that v_1 is about 25 km./hr. From figure 12 a reasonable value for $\nabla N_m(S)$ is 0·30, corresponding to a variation of about $\pm$ 1·5 Mcyc./sec. This quantity is, of course, difficult to estimate, because of the presence of the much larger variation due to the solar ionizing influence, but the figure suggested would seem to be the measure of the daytime departure from the theoretical Chapman diurnal variation This gives $\sigma_1 v_1/\omega_1 = 0{\cdot}3$ and $\sigma_1 = 0{\cdot}006$ radians/km. The relative phases of the $z_m(S)$ and $N_m(S)$ variations show that σ_1 is positive in sign. This is confirmed also by the steady overall *downward* movement of the region (cf. figure 14).

The application of equation (11) to the data in figure 10 gives $\sigma_1 v_2/\omega_1$ as about 0 08 in the daytime, so that v_2 is about 6 6 km./hr This value of v_2 is also consistent with the amplitude of $\nabla z_m(LS)$ in the daytime.

It is clear that it will not be easy to interpret completely the rather complex variations in figures 9 and 10. In the first place, $\sigma_1 v_1/\omega_1$ is not much less than unity, so that the simple results of equations (5) and (6) are inapplicable Secondly, the lunar velocity v_2 is not much less than the solar v_1, so that the applicability of the treatment in (5·2·1) is doubtful. These circumstances arise, of course, from the particularly large amplitude of the solar variations in the F_2 region at Huancayo, together with the particularly large ratio of the lunar to the solar variation at that place. (It is worthy of remark that these two features are also prominent in the solar and lunar magnetic variations at Huancayo) It is possible, however, to interpret some of the major effects in terms of the theory, although a more detailed application of the latter will be reserved for the data from observatories at higher latitudes, which give smaller, but simpler, lunar-solar variations.

It is observed from comparison of figures 5 and 8 that the average lunar variations in z_m and N_m are approximately in phase quadrature, the latter leading the former by about 3·5 hr It is not easy to explain this result, which is a feature of the results from all locations so far analyzed, save in terms of the theory developed in (5·2) If it were possible to ignore the disturbing effect of the large S variation on the non-linear system represented by equations (5) and (6), then the conclusion could be drawn that, since ∇N_m leads ∇z_m, σ_2 is negative in sign. This is probably not a safe deduction however at Huancayo, in view of the obviously large interaction between the L and S variations revealed by figures 9 and 10. In general, the variation in N_m should show simpler features than that in z_m. It was shown in I that while N_m was a simple function of the velocity-height gradient, the variation of z_m with this parameter was more complicated, and was liable to change of phase at critical values of the latter. Comparison of figures 9 and 10 shows that the N_m variation is considerably simpler than that in z_m, as might be anticipated from such reasoning.

Moreover in the hours from midnight to dawn the F region has sunk (figure 12) to a level where there is strong reason to believe that attachment and recombination is important (The F_1 region at 200 km. disappears very rapidly at sunset.) It is likely that the steadiness of the F region at the height of about 290 km. at this time is simply due to the fact that electrons moving downwards below this level disappear rapidly by these processes $\nabla z_m(LS)$ may then be expected to be small at this time, as figure 9 shows to be the case. On the other hand the N_m variation should remain, and could even be amplified, for in this case the vertical lunar velocity will add to, or subtract from, the solar velocity which is carrying electrons downwards into the recombination 'sink'.

The most striking feature of $\nabla N_m(LS)$ is the continuous retardation of phase which takes place through most of the solar day The average rate of this retardation, which becomes prominent soon after noon, when the F_2 region is fully formed, and ion production is diminishing, is 1 hr./hr. It continues until dawn, when the new F_2 region is formed and the phase changes abruptly This phase retardation is a necessary consequence of equation (11), which explains it in both magnitude and sense. The smallness of $\nabla N_m(LS)$ at 23 hr probably indicates that $\sin S$ is passing through a zero near that time. If such be the case then equation (11) shows that $\nabla N_m(LS)$ should change phase by π (or 6 hr.) at this time. A rapid variation of this amount, superimposed on the steady retardation of phase, is seen to occur then (figure 10).

Summing up, it appears that the theory outlined in §§ 5·2 and 5·2 1 can explain the main features of the solar and lunar-solar variations in N_m and z_m It is not adequate in its present form, however, to explain the details of these variations in view of the large amplitude of both variations, and the consequent high degree of interaction present

This work has been carried out on behalf of the Australian Council for Scientific and Industrial Research The author is indebted to the Commonwealth Astronomer, Dr R. v. d. R. Woolley, for providing facilities and computers, and for useful discussion, also to Dr C W Allen and Mr G Odgers of the Observatory Staff Mr C B. Kirkpatrick of the Radio Research Board's staff gave valuable mathematical assistance The very considerable burden of computing has been shared by many persons, notably Misses J Foley and H. Champion, to whom the author's thanks are specially due All the data used are published by kind permission of the Director of the Department of Terrestrial Magnetism and Electricity, Carnegie Institution of Washington

References

Appleton & Weekes 1939 *Proc. Roy Soc.* A, 171, 171
Chapman & Bartels 1940 *Geomagnetism* Oxford Clarendon Press.
Martyn 1947 *Proc. Roy. Soc.* A, 189, 241.
Pekeris 1937 *Proc. Roy Soc.* A, 158, 650.

The theory of the magneto-resistance effects in metals

By E. H. Sondheimer, B A., and A. H. Wilson, F.R S.

Trinity College, Cambridge

(*Received* 14 *October* 1946)

General formulae are obtained for the effect of a magnetic field on the electrical and thermal conductivities of a metal in which there are two overlapping bands of normal form Simple formulae are set up which, though not strictly valid for all temperatures and fields, reduce to the correct expressions in the three limiting cases of high temperatures, low temperatures and very strong magnetic fields

The behaviour of the electrical resistance at low temperatures is discussed, and it is shown that in certain cases the resistance may pass through a minimum as the temperature is increased provided the magnetic field is large enough It is also shown that in general the Lorenz number is increased by the presence of a magnetic field, but that the thermal conductivity of the lattice is unaffected by a magnetic field

1. Introduction

While the modern theory of metals gives a satisfactory account of most of the conduction phenomena, there are still many problems concerning the effect of a magnetic field which have not been adequately explained. In all cases it is possible to give formal expressions for the physical quantities concerned, but comparison with experiment is impossible unless a sufficiently simple model is adopted which enables the calculations to be completed It is usual to assume that the electrons are quasi-free, the energy being proportional to the square of the wave-number, and, even if less restrictive assumptions are made in parts of the calculation, the assumption of quasi-free electrons is generally made at the end Unfortunately, this simple model cannot be used for discussing the magneto-resistance effects since it gives a zero change of resistance, and more complicated models are necessary (Peierls 1931). One assumption that enables a definite result to be obtained is that the energy is a general quadratic function of the wave-vector and that a time of relaxation exists.* The disadvantage of this model is that a time of relaxation can only be defined at high temperatures even for the simplest model, and that in the

* The time of relaxation is defined by the rate at which a non-equilibrium distribution of the electrons is brought into the equilibrium state when the electrons are subject only to collisions with the lattice. In general a time of relaxation cannot be unambiguously defined since it depends upon the initial distribution. For an isotropic substance, however, a time of relaxation exists at temperatures large compared with the Debye temperature, and all the calculations then become relatively simple. At low temperatures, on the other hand, this simplifying assumption is no longer permissible (Wilson 1936, pp. 159, 208, 212, 219)

One of the consequences of the existence of a time of relaxation is the validity of the Wiedemann-Franz law (see Wilson 1936, p. 175), and more particularly Makinson (1938), where a detailed discussion is given of the deviations from the Wiedemann-Franz law).

model proposed it is not possible to calculate the time of relaxation in terms of the atomic constants of the metal The calculation is therefore essentially a formal one.

There are only two models for which it has so far proved possible to carry out the calculations completely. The first refers to semi-conductors in which the electrons are considered to be free and to have a Maxwellian distribution (Harding 1933); it will not be considered further in the present paper. The second postulates the existence of two overlapping bands in each of which the energy is proportional to the square of the wave-number This model is somewhat artificial, but it is the only one which is sufficiently simple to be tractable The main object of the present paper is to see how far the predictions of this model can be used to interpret the experimental facts, particularly at low temperatures where the use of a time of relaxation is not admissible. Only the electrical and thermal conductivities are discussed, the thermo-magnetic effects being left for a subsequent paper Further, the cases in which the magnetic field is parallel to the electrical or thermal current are not considered, since the present model is not sufficiently general to give a non-zero result in such cases

In § 2 we collect together the general formulae required for discussing the two-band model, and in § 3 the calculations are completed for the three limiting cases of very high and very low temperatures and very strong magnetic fields The comparison with the experimental results for the electrical and thermal conductivities is given in §§ 4 and 6.

The electrical resistance in a magnetic field of arbitrary magnitude has been studied by many authors, and in particular by Bronstein (1932) and Jones (1936) in the relatively simple case of high temperatures where it is legitimate to assume that a time of relaxation exists. We show here that the same formulae can be used over the whole temperature range if certain approximations are made and if the various quantities which occur in the formulae are suitably generalized To do this we obtain exact solutions of the transport equation in three limiting cases, namely, high temperatures, low temperatures and very strong magnetic fields. We then show that a comparatively simple formula (equation (57)) reduces to the correct form in these three limiting cases, though it cannot be valid for all temperatures and all fields We further set up a corresponding equation for the effect of a magnetic field on the electronic thermal conductivity (equation (74))

Having established the approximate validity of the formulae at low temperatures we use them to explain some results of Milner (1937) which have so far proved difficult to fit into their proper place In experiments on cadmium at liquid helium temperatures, Milner found that in general the resistance increases with the temperature for a given magnetic field For sufficiently large fields, however, the resistance first diminishes as the temperature is increased from the lowest available value; it then passes through a minimum and finally increases. We show in § 4·3 that in certain circumstances the resistance of the two-band model behaves in just this manner.

In § 5 we discuss Matthiessen's rule, and we show that the strict additivity of the residual and ideal resistances is not to be expected except in special cases

The electronic thermal conductivity is considered in § 6, and formulae are given for the effect of a magnetic field on the Lorenz number It is shown that in general the Lorenz number is increased by the presence of a magnetic field, which is in agreement with the experimental facts.

In § 7 we consider the thermal conductivity due to the lattice, and we show that it is unaffected by the presence of a magnetic field This gives a partial justification for the experimental procedure of determining the lattice thermal conductivity by applying such a strong magnetic field that the electronic thermal conductivity is reduced to negligible proportions Since, however, most models show that the electrical and thermal resistivities should saturate in very strong magnetic fields and not become infinite, caution must be used when extrapolating the results obtained with finite fields

2. General formulae

We use the well-known model in which there are two overlapping partially filled bands of normal form, the s- and d-bands We assume that the s-d transitions are negligible. This restriction can readily be removed (see Wilson 1938), but nothing is gained thereby, as it merely complicates the result without introducing any phenomena which have not been fully discussed before With this simplifying assumption the behaviour of the electrons in each band is independent of what happens in the other band, and the contributions of the bands to the currents can be calculated separately We also neglect the quantization of the electron orbits due to the magnetic field, an effect which has been considered by Titeica (1935), which means that our results require a correction when the temperature is very low or the magnetic field is very high. We hope to consider the magnitude of this correction in a subsequent paper

We use the standard notation (see, for example, Wilson 1936), but supplement it by inserting a suffix s or d, referring to the s- and d-band, when necessary to avoid confusion.

The energy levels in the s-band are given by

$$E = \frac{\hbar^2 |\mathbf{k}|^2}{2m_s}, \tag{1}$$

where $\mathbf{k}$ is the wave-vector with components k_1, k_2, k_3, and m_s is the effective mass of the electrons, while the corresponding expression for the energy levels in the d-band is given by

$$E = A - \frac{\hbar^2 |\mathbf{k}|^2}{2m_d}, \tag{2}$$

where A is the energy overlap of the two bands Since both bands are supposed to be partially occupied by electrons, the Fermi energy level ζ must be such that

$0 < \zeta < A$ The distribution function f for the electrons is most conveniently written as

$$f = f_0 - \{k_1 c_1(E) + k_2 c_2(E)\} \frac{\partial f_0}{\partial E}, \tag{3}$$

where $c_1(E)$ and $c_2(E)$ are two functions of the energy which have to be determined, and where f_0 is the Fermi function,*

$$f_0 = \frac{1}{e^{(E-\zeta)/kT} + 1},$$

there being two separate distribution functions, one for each band (When considering the d-electrons it is convenient to change the signs of k_1 and k_2 in equations (3) and (4) so as to ensure that equations (7) and (8) apply to both the s- and the d-electrons. This is the convention adopted here, but any consistent convention can be used) The expression (3) is the correct form for f when there is a magnetic field $(0, 0, H)$, an electric field $(\mathscr{E}_x, \mathscr{E}_y, 0)$ and a temperature gradient $(\partial T/\partial x, \partial T/\partial y, 0)$, which is sufficiently general for our present purpose.

In the steady state, f is determined by the Boltzmann equation

$$\left[\frac{\partial f}{\partial t}\right]_{\text{fields}} + \left[\frac{\partial f}{\partial t}\right]_{\text{coll}} = 0. \tag{4}$$

The general expression for $[\partial f/\partial t]_{\text{fields}}$ is (Wilson 1936, p 158, equation (226))

$$\left[\frac{\partial f}{\partial t}\right]_{\text{fields}} = \frac{\epsilon}{\hbar}\left(\mathscr{E} + \frac{1}{c}\mathbf{v} \times \mathbf{H}\right) \operatorname{grad}_{\mathbf{k}} f - \mathbf{v} \operatorname{grad}_{\mathbf{r}} f,$$

where $-\epsilon$ is the electronic charge, $\mathbf{v}$ is the velocity of an electron ($\hbar\mathbf{v} = \operatorname{grad}_{\mathbf{k}} E$), c is the velocity of light and f is supposed to be a function of the wave-vector $\mathbf{k}$ and the space-vector $\mathbf{r}$ Now f_0 only involves the space coordinates through T and ζ, which in a homogeneous metal is a function of the temperature only. Hence, if we insert (3) into the above expression for $[\partial f/\partial t]_{\text{fields}}$ and neglect as usual products of $c_1(E)$ or $c_2(E)$ with $\mathscr{E}$ or grad T but retain all the terms involving H, we find for the s-electrons

$$\left[\frac{\partial f}{\partial t}\right]_{\text{fields}} = \frac{\hbar}{m}\frac{\partial f_0}{\partial E}\left[k_1\left\{\epsilon\mathscr{E}_x + T\frac{\partial}{\partial x}\left(\frac{\zeta}{T}\right) + \frac{E}{T}\frac{\partial T}{\partial x} + \frac{\epsilon H}{\hbar c}c_2(E)\right\}\right.$$
$$\left. + k_2\left\{\epsilon\mathscr{E}_y + T\frac{\partial}{\partial y}\left(\frac{\zeta}{T}\right) + \frac{E}{T}\frac{\partial T}{\partial y} - \frac{\epsilon H}{\hbar c}c_1(E)\right\}\right], \tag{5}$$

the suffix s being omitted $[\partial f/\partial t]_{\text{coll}}$ is the rate of change of f due to collisions between the electrons and the lattice Its explicit form is given in § 3, but for the present it is only necessary to know that it involves the unknown functions $c_1(E)$ and $c_2(E)$ linearly, so that we may write for the s-electrons

$$\left[\frac{\partial f}{\partial t}\right]_{\text{coll.}} = -\frac{\hbar}{m}\{k_1 L(c_1) + k_2 L(c_2)\}\frac{\partial f_0}{\partial E}, \tag{6}$$

* The use of k in two different senses, namely, k as Boltzmann's constant and $\mathbf{k}$ as the wave-vector with the components k_1, k_2, k_3, should not cause any confusion Similarly, c is used to denote the velocity of light, while $c_1(E)$ and $c_2(E)$ are distribution functions.

where L is an integral operator and the suffix s has been again omitted Combining (4), (5) and (6) we obtain the following simultaneous equations for c_1 and c_2

$$L(c_1) - \frac{\epsilon H}{\hbar c} c_2 = \epsilon \mathscr{E}_x + T \frac{\partial}{\partial x}\left(\frac{\zeta}{T}\right) + \frac{E}{T}\frac{\partial T}{\partial x}, \tag{7}$$

$$L(c_2) + \frac{\epsilon H}{\hbar c} c_1 = \epsilon \mathscr{E}_y + T \frac{\partial}{\partial y}\left(\frac{\zeta}{T}\right) + \frac{E}{T}\frac{\partial T}{\partial y}. \tag{8}$$

When these equations have been solved, the electric current $(J_x, J_y, 0)$ and the heat current $(w_x, w_y, 0)$ can be written down For example, we have

$$J_{xs} = -\frac{\epsilon}{4\pi^3}\iiint v_{xs} f_s \, dk_1 dk_2 dk_3$$

Since $\hbar \mathbf{v} = \operatorname{grad}_{\mathbf{k}} E$, this becomes

$$J_{xs} = \frac{(2m_s)^{\frac{1}{2}}\epsilon}{3\pi^2\hbar^4}\int E^{\frac{1}{2}} c_{1s}(E) \frac{\partial f_0}{\partial E} dE, \tag{9}$$

on changing to polar coordinates in the $\mathbf{k}$ space and using E instead of $|\mathbf{k}|$ as the final variable of integration. Similarly,

$$w_{xs} = \frac{1}{4\pi^3}\int v_{xs} E f_s \, dk_1 dk_2 dk_3 = -\frac{(2m_s)^{\frac{1}{2}}}{3\pi^2\hbar^4}\int E^{\frac{1}{2}} c_{1s}(E) \frac{\partial f_0}{\partial E} dE \tag{10}$$

There are corresponding expressions for J_y and w_y with c_1 replaced by c_2

In writing down the currents for the d-band it must be borne in mind that $\mathbf{k} = 0$ does not correspond to $E = 0$, as it does for the s-band, but to $E = A$. The expressions corresponding to (9) and (10) are therefore

$$J_{xd} = \frac{(2m_d)^{\frac{1}{2}}\epsilon}{3\pi^2\hbar^4}\int E'^{\frac{1}{2}} c_{1d}(E') \frac{\partial f_0}{\partial E'} dE', \tag{11}$$

$$w_{xd} = -\frac{(2m_d)^{\frac{1}{2}}}{3\pi^2\hbar^4}\int E E'^{\frac{1}{2}} c_{1d}(E') \frac{\partial f_0}{\partial E'} dE', \tag{12}$$

where $E' = A - E$.

The total electrical and thermal currents are the sums of those due to the two bands separately, and it can readily be shown that they are of the form

$$J_x = \frac{\epsilon}{3\pi^2\hbar^4}\left[(i_{1s} + i_{1d})\left\{\epsilon \mathscr{E}_x + T\frac{\partial}{\partial x}\left(\frac{\zeta}{T}\right)\right\} + (i_{2s} - i_{2d} + A i_{1d})\frac{1}{T}\frac{\partial T}{\partial x}\right.$$
$$\left. - (i_{4s} - i_{4d})\left\{\epsilon \mathscr{E}_y + T\frac{\partial}{\partial y}\left(\frac{\zeta}{T}\right)\right\} - (i_{5s} + i_{5d} - A i_{4d})\frac{1}{T}\frac{\partial T}{\partial y}\right], \tag{13}$$

$$J_y = \frac{\epsilon}{3\pi^2\hbar^4}\left[(i_{4s} - i_{4d})\left\{\epsilon \mathscr{E}_x + T\frac{\partial}{\partial x}\left(\frac{\zeta}{T}\right)\right\} + (i_{5s} + i_{5d} - A i_{4d})\frac{1}{T}\frac{\partial T}{\partial x}\right.$$
$$\left. + (i_{1s} + i_{1d})\left\{\epsilon \mathscr{E}_y + T\frac{\partial}{\partial y}\left(\frac{\zeta}{T}\right)\right\} + (i_{2s} - i_{2d} + A i_{1d})\frac{1}{T}\frac{\partial T}{\partial y}\right], \tag{14}$$

$$w_x = -\frac{1}{3\pi^2\hbar^4}\Big[(\imath_{2s} - i_{2d} + A\imath_{1d})\Big(\epsilon\mathscr{E}_x + T\frac{\partial}{\partial x}\Big(\frac{\zeta}{T}\Big)\Big) + (i_{3s} + i_{3d} - 2Ai_{2d} + A^2\imath_{1d})\frac{1}{T}\frac{\partial T}{\partial x}$$
$$-(\imath_{5s} + \imath_{5d} - A\imath_{4d})\Big(\epsilon\mathscr{E}_y + T\frac{\partial}{\partial y}\Big(\frac{\zeta}{T}\Big)\Big) - (i_{6s} - \imath_{6d} + 2A\imath_{5d} - A^2\imath_{4d})\frac{1}{T}\frac{\partial T}{\partial y}\Big], \quad (15)$$

$$w_y = -\frac{1}{3\pi^2\hbar^4}\Big[(\imath_{5s} + \imath_{5d} - A\imath_{4d})\Big(\epsilon\mathscr{E}_x + T\frac{\partial}{\partial x}\Big(\frac{\zeta}{T}\Big)\Big) + (\imath_{6s} - \imath_{6d} + 2A\imath_{5d} - A^2\imath_{4d})\frac{1}{T}\frac{\partial T}{\partial x}$$
$$+(\imath_{2s} - \imath_{2d} + A\imath_{1d})\Big(\epsilon\mathscr{E}_y + T\frac{\partial}{\partial y}\Big(\frac{\zeta}{T}\Big)\Big) + (\imath_{3s} + \imath_{3d} - 2Ai_{2d} + A^2\imath_{1d})\frac{1}{T}\frac{\partial T}{\partial y}\Big]. \quad (16)$$

In the above equations $\imath_s$ is a quantity depending upon the magnetic field and the s-electron variables, while i_d is the corresponding quantity for the d-electrons. In order to show that the currents have the forms given above, the simultaneous integral equations (7) and (8) for c_1 and c_2 have to be solved formally in terms of the resolving nuclei, and the resulting expressions for c_1 and c_2 inserted into equations (9) to (12). In this way we obtain formal expressions for the coefficients that occur in equations (13) to (16), from which the symmetry relations between the coefficients can be verified at once The proof is straightforward but cumbersome and is left to the reader. So far as the general theory is concerned, all that is required is the general form of the equations, the symmetry relations between the coefficients, and the relations (18) No further progress can be made unless the integral equations can be solved explicitly. This is done in §§ 3 1 and 3·2 for two special cases, and explicit expressions for the $\imath$'s are given there (It is obvious by inspection of equations (7) and (8) that the currents must involve the electric field and temperature gradient linearly, and that they must therefore have the general form given in equations (13) to (16), but the relations between the coefficients cannot be determined except by formally solving the integral equations The symmetry relations in the most general case have been determined by Kohler (1941), and the equations given above are particular cases of his)

To proceed further the integrals $\imath_s$ and i_d must be evaluated, which is done by the usual approximate method for integrals involving the Fermi distribution function. First, i_s is expanded as an ascending series in $(kT/\zeta)^2$ and $\imath_d$ as an ascending series in $(kT/\zeta')^2$, where $\zeta' = A - \zeta$

$$i_s = \sum_n (kT/\zeta)^{2n}\imath_s^{(n)}, \quad i_d = \sum_n (kT/\zeta')^{2n}\imath_d^{(n)} \quad (17)$$

All the physical quantities of interest can then be expressed in terms of the $\imath^{(n)}$'s, but for ordinary temperatures (which are well below the temperature at which the electron gas becomes non-degenerate) it is sufficient to retain only the lowest $\imath^{(n)}$'s which give a non-zero result. In deciding what approximation to go to, the following relations are required:

$$i_{2s}^{(0)} = \zeta i_{1s}^{(0)}, \quad i_{3s}^{(0)} = \zeta^2\imath_{1s}^{(0)}, \quad i_{5s}^{(0)} = \zeta i_{4s}^{(0)}, \quad i_{6s}^{(0)} = \zeta^2 i_{4s}^{(0)}, \quad (18)$$

together with the corresponding relations for the $\imath_d$'s with ζ replaced by ζ'. These relations are easily proved in general by inserting the formal expressions for c_1

and c_2 into the integrals and evaluating the ι's to the zero order by replacing E everywhere by ζ The proof is left to the reader, the relations can also be verified in the particular cases of interest here from the explicit expressions given in § 3.

The electrical conductivity σ is measured by applying an electric field in say the x direction and observing the current in this direction. No electric current is allowed to flow in the transverse direction and the temperature of the specimen is kept uniform, so that $J_y = 0$ and $\partial T/\partial x = \partial T/\partial y = 0$ Equation (14) determines the transverse electric field $\mathscr{E}_y$ set up, and, if $\mathscr{E}_y$ is eliminated, equation (13) gives

$$\sigma = \frac{J_x}{\mathscr{E}_x} = \frac{e^2}{3\pi^2\hbar^4}\frac{(V_s+V_d)^2+(W_s-W_d)^2}{V_s+V_d}, \tag{19}$$

where

$$V = \iota_1^{(0)}, \quad W = \iota_4^{(0)} \tag{20}$$

The conditions under which κ, the thermal conductivity, is measured do not seem to have been standardized No electric current flows, so that $J_x = J_y = 0$ and $\partial T/\partial x \neq 0$, but the thermal conditions in the y direction are a matter of some speculation In general, it would seem that the experimental conditions are such that, approximately, $w_y = 0$, since the specimen is usually in the form of a long strip whose ends are kept at a constant difference of temperature while no attempt is made to control the temperature of the edges, which are only cooled by radiation It will therefore be assumed that the correct condition is $w_y = 0$ Then

$$\kappa = \frac{w_x}{-\partial T/\partial x} = \frac{k^2}{3\pi^2\hbar^4}T\frac{(X_s+X_d)^2+(Y_s-Y_d)^2}{X_s+X_d}, \tag{21}$$

where

$$\zeta^2 X = \zeta^2\iota_1^{(1)} - 2\zeta\iota_2^{(1)} + \iota_3^{(1)}, \quad \zeta^2 Y = \zeta^2\iota_4^{(1)} - 2\zeta\iota_5^{(1)} + \iota_6^{(1)}. \tag{22}$$

The expressions for all the other galvano- and thermo-magnetic effects can be deduced from equations (13)–(16)

3 Special solutions

3 1. *High temperatures* At high temperatures such that $(\Theta/T)^2$ can be neglected, where Θ is the Debye temperature, a time of relaxation exists It is therefore permissible to write (Wilson 1936, pp. 159, 208)

$$\left[\frac{\partial f}{\partial t}\right]_{\text{coll}} = -\frac{f-f_0}{\tau}, \tag{23}$$

where the time of relaxation τ is in general different for the two bands. Equation (23) should therefore be applied to the two bands separately, and τ_s and τ_d are functions of the energy only. Omitting the suffixes s and d, and combining equations (3), (4), (5) and (23), the following equations for c_1 and c_2 are obtained

$$\left.\begin{aligned}\frac{m}{\hbar\tau}c_1 + \alpha c_2 &= -\left(e\mathscr{E}_x + T\frac{\partial}{\partial x}\left(\frac{\zeta}{T}\right) + \frac{E}{T}\frac{\partial T}{\partial x}\right),\\ -\alpha c_1 + \frac{m}{\hbar\tau}c_2 &= -\left(e\mathscr{E}_y + T\frac{\partial}{\partial y}\left(\frac{\zeta}{T}\right) + \frac{E}{T}\frac{\partial T}{\partial y}\right),\end{aligned}\right\} \tag{24}$$

where $\alpha = eH/\hbar c$ It is easy to verify that the currents have the form shown in equations (13) to (16), and the explicit expressions for the i's are found to be

$$i_n = -(2m)^{\frac{1}{2}}\frac{m}{\hbar}\int \frac{E^{n+\frac{1}{2}}}{\tau\{\alpha^2+(m/\hbar\tau)^2\}}\frac{\partial f_0}{\partial E}dE \quad (n = 1, 2, 3), \tag{25}$$

$$i_n = -(2m)^{\frac{1}{2}}\alpha\int \frac{E^{n-\frac{1}{2}}}{\alpha^2+(m/\hbar\tau)^2}\frac{\partial f_0}{\partial E}dE \quad (n = 4, 5, 6). \tag{26}$$

Now expand the i's in powers of $(kT/\zeta)^2$ by means of the well-known formula

$$-\int \phi(E)\frac{\partial f_0}{\partial E}dE = \phi(\zeta) + \tfrac{1}{6}(\pi kT)^2\frac{d^2\phi}{d\zeta^2}.$$

This gives

$$V = \frac{m}{\hbar\tau}\frac{(2m\zeta)^{\frac{1}{2}}}{\alpha^2+(m/\hbar\tau)^2}, \quad W = \alpha\frac{(2m\zeta)^{\frac{1}{2}}}{\alpha^2+(m/\hbar\tau)^2}, \tag{27}$$

$$X = \tfrac{1}{3}\pi^2 V, \quad Y = \tfrac{1}{3}\pi^2 W \tag{28}$$

3·2. *Low temperatures.* At sufficiently low temperatures where the residual resistance predominates the method of successive approximations given by Wilson (1937) can be used When scattering by both impurities and the lattice vibrations is considered, the collision operator L is given by (Wilson 1937, equation (19)*)

$$L(c) = -\beta E^{\frac{1}{2}}c + E^{-\frac{1}{2}}L'(c). \tag{29}$$

Here β is a constant determining the effect of the impurities (in Wilson's notation $\beta = (2m)^{\frac{1}{2}}M/4\hbar^3$) and

$$L'(c) = -\frac{(2m)^{\frac{1}{2}}}{4\hbar^3\Lambda}\left(\frac{T}{\Theta}\right)^3 \times \int_{-\Theta/T}^{\Theta/T}\left[Ec(E) - c(E+kTz)\left\{E + \tfrac{1}{2}kTz - D\left(\frac{T}{\Theta}\right)^2 z^2\right\}\right]\frac{e^{\eta}+1}{e^{\eta+z}+1}\frac{z^2dz}{|1-e^{-z}|}. \tag{30}$$

The usual notation is employed here, with $\eta = (E-\zeta)/kT$, and Λ and D are constants characteristic of the pure metal

Now solve the equations (7) and (8) by successive approximations, first neglecting $L'(c)$ in order to obtain the zero approximation $c^{(0)}$. Next substitute $c = c^{(0)} + c^{(1)}$ and neglect $L'(c^{(1)})$ in order to determine $c^{(1)}$, and so on. The zero and first-order approximations then give the following expressions for the i's

$$i_1 = -(2m)^{\frac{1}{2}}\beta\int \frac{E^2}{\alpha^2+\beta^2E}\frac{\partial f_0}{\partial \eta}d\eta + (2m)^{\frac{1}{2}}(\alpha^2K_{0,0} - \beta^2K_{\frac{1}{2},\frac{1}{2}}), \tag{31}$$

$$i_2 = -(2m)^{\frac{1}{2}}\beta\int \frac{E^2}{\alpha^2+\beta^2E}\frac{\partial f_0}{\partial \eta}d\eta + (2m)^{\frac{1}{2}}(\alpha^2K_{1,0} - \beta^2K_{\frac{3}{2},\frac{1}{2}}), \tag{32}$$

* The expression on the right of this equation is not $L(c)$ but a multiple of it The multiplying constant is most easily determined by an inspection of equations (3) and (19) of Wilson's paper and equations (4), (5) and (6) of the present paper.

$$\iota_3 = -(2m)^{\frac{1}{2}}\beta\int\frac{E^4}{\alpha^2+\beta^2E}\frac{\partial f_0}{\partial\eta}d\eta + (2m)^{\frac{1}{2}}(\alpha^2K_{1,1}-\beta^2K_{\frac{1}{2},\frac{1}{2}}), \tag{33}$$

$$\iota_4 = -(2m)^{\frac{1}{2}}\alpha\int\frac{E^{\frac{1}{2}}}{\alpha^2+\beta^2E}\frac{\partial f_0}{\partial\eta}d\eta - (2m)^{\frac{1}{2}}\alpha\beta(K_{0,\frac{1}{2}}+K_{\frac{1}{2},0}), \tag{34}$$

$$\iota_5 = -(2m)^{\frac{1}{2}}\alpha\int\frac{E^{\frac{1}{2}}}{\alpha^2+\beta^2E}\frac{\partial f_0}{\partial\eta}d\eta - (2m)^{\frac{1}{2}}\alpha\beta(K_{1,\frac{1}{2}}+K_{\frac{1}{2},0}), \tag{35}$$

$$\iota_6 = -(2m)^{\frac{1}{2}}\alpha\int\frac{E^{\frac{1}{2}}}{\alpha^2+\beta^2E}\frac{\partial f_0}{\partial\eta}d\eta - (2m)^{\frac{1}{2}}\alpha\beta(K_{1,\frac{1}{2}}+K_{\frac{1}{2},1}), \tag{36}$$

where
$$K_{m,n} = \int\frac{E^m}{\alpha^2+\beta^2E}L'\left(\frac{E^n}{\alpha^2+\beta^2E}\right)\frac{\partial f_0}{\partial\eta}d\eta = K_{n,m} \tag{37}$$

The simplest way to obtain V, W, X, Y explicitly is to proceed as follows By use of equation (11) of Wilson's paper, $K_{m,n}$ may be put into the form

$$K_{m,n} = -\frac{(2m)^{\frac{1}{2}}}{4\hbar^3\Lambda}\left(\frac{T}{\Theta}\right)^3\int_{-\Theta/T}^{\Theta/T}\frac{z^2dz}{(e^z-1)\,|\,1-e^{-z}\,|} \times\int_{-\infty}^{\infty}\{G_{m,n}(E,z)-G_{m,n}(E-kTz,z)\}\frac{\partial f_0}{\partial\eta}d\eta, \tag{38}$$

where
$$G_{m,n}(E,z) = \frac{1}{kT}\int^E F_{m,n}(E',z)\,dE' \tag{39}$$

and
$$F_{m,n}(E,z) = \frac{E^{m+n+1}}{(\alpha^2+\beta^2E)^2} - \frac{E^m(E+kTz)^n}{(\alpha^2+\beta^2E)\{\alpha^2+\beta^2(E+kTz)\}}\left\{E+\tfrac{1}{2}kTz-D\left(\frac{T}{\Theta}\right)^2z^2\right\} \tag{40}$$

Now expand $F_{m,n}(E,z)$ in powers of kTz, but treat the term $D(T/\Theta)^2z^2$ as a constant, so that

$$F_{m,n}(E,z) = A_{m,n}(E,z)+kTzB_{m,n}(E,z)+(kTz)^2C_{m,n}(E,z)+\quad,$$

where A, B, C, . are even functions of z, and in particular

$$A_{m,n}(E,z) = D\left(\frac{T}{\Theta}\right)^2z^2\frac{E^{m+n}}{(\alpha^2+\beta^2E)^2}$$

Then
$$G_{m,n}(E,z)-G_{m,n}(E-kTz,z)-G_{m,n}(E,-z)+G_{m,n}(E+kTz,-z) = 2zA_{m,n}+k^2T^2z^3(\tfrac{1}{3}A''_{m,n}-B'_{m,n}+2C_{m,n})+\quad.,$$

where dashes denote differentiation with respect to E Substitution of this into (38) and evaluation of the integrals to the second order in kT/ζ then gives

$$K_{m,n} = \frac{(2m)^{\frac{1}{2}}}{4\hbar^3\Lambda}\left(\frac{T}{\Theta}\right)^3\left[2\mathscr{I}_0D\left(\frac{T}{\Theta}\right)^2\frac{\zeta^{m+n}}{(\alpha^2+\beta^2\zeta)^2} + k^2T^2\int_0^{\Theta/T}\{\tfrac{1}{3}\pi^2A''_{m,n}+z^2(\tfrac{1}{3}A''_{m,n}-B'_{m,n}+2C_{m,n})\}\frac{z^3dz}{(e^z-1)(1-e^{-z})}\right], \tag{41}$$

where $A, B, C, \ldots$ are now functions of ζ and z, and dashes denote differentiation with respect to ζ. Also

$$\mathscr{J}_n = \int_0^{\Theta/T} \frac{z^n dz}{(e^z-1)(1-e^{-z})}.$$

In this way expressions for V and W are at once obtained. To find X explicitly expressions like the following are needed.

$$\zeta^2(\imath_1^{(0)})'' - 2\zeta(\zeta\imath_1^{(0)})'' + (\zeta^2\imath_1^{(0)})'' = 2\imath_1^{(0)} = 2V,$$

$$\zeta^2 A''_{0,0} - \zeta A''_{1,0} - \zeta A''_{0,1} + A''_{1,1} = 2A_{0,0},$$

and

$$\zeta^2 B'_{0,0} - \zeta B'_{1,0} - \zeta B'_{0,1} + B'_{1,1},$$

which is the coefficient of kTz in $-\zeta F_{0,0} + F_{0,1}$, and is

$$-\left(\zeta - D\left(\frac{T}{\Theta}\right)^2 z^2\right)\frac{1}{(\alpha^2+\beta^2\zeta)^2}.$$

When all the terms are collected together and simplified, the final results are

$$V = (2m\zeta)^{\frac{1}{2}}\frac{\beta\zeta^{\frac{1}{2}}}{\alpha^2+\beta^2\zeta} + 4\left(\frac{T}{\Theta}\right)^5 \mathscr{J}_5 \frac{m^3 D}{\hbar^3 \Lambda}\frac{\alpha^2-\beta^2\zeta}{(\alpha^2+\beta^2\zeta)^2}, \tag{42}$$

$$W = (2m\zeta)^{\frac{1}{2}}\frac{\alpha}{\alpha^2+\beta^2\zeta} - 4\left(\frac{T}{\Theta}\right)^5 \mathscr{J}_5 \frac{m^3 D}{\hbar^3 \Lambda}\frac{2\alpha\beta\zeta^{\frac{1}{2}}}{(\alpha^2+\beta^2\zeta)^2}, \tag{43}$$

$$X = \tfrac{1}{3}\pi^2(2m\zeta)^{\frac{1}{2}}\frac{\beta\zeta^{\frac{1}{2}}}{\alpha^2+\beta^2\zeta} + 2\left(\frac{T}{\Theta}\right)^3\left\{\mathscr{J}_5 + \frac{D}{\zeta}\left(\frac{T}{\Theta}\right)^2\left(\frac{2\pi^2}{3}\mathscr{J}_5 - \tfrac{1}{3}\mathscr{J}_7\right)\right\}\frac{m^3\zeta}{\hbar^3\Lambda}\frac{\alpha^2-\beta^2\zeta}{(\alpha^2+\beta^2\zeta)^2}, \tag{44}$$

$$Y = \tfrac{1}{3}\pi^2(2m\zeta)^{\frac{1}{2}}\frac{\alpha}{\alpha^2+\beta^2\zeta} - 2\left(\frac{T}{\Theta}\right)^3\left\{\mathscr{J}_5 + \frac{D}{\zeta}\left(\frac{T}{\Theta}\right)^2\left(\frac{2\pi^2}{3}\mathscr{J}_5 - \tfrac{1}{3}\mathscr{J}_7\right)\right\}\frac{m^3\zeta}{\hbar^3\Lambda}\frac{2\alpha\beta\zeta^{\frac{1}{2}}}{(\alpha^2+\beta^2\zeta)^2}. \tag{45}$$

3·3 *Strong magnetic fields* Kohler (1940) has given a method of successive approximations which is useful when the magnetic field is very large. The zero-order approximation is obtained by neglecting entirely all the terms arising from $[\partial f/\partial t]_{\text{coll}}$. The higher approximations can be obtained by proceeding exactly as in the preceding section. After a considerable amount of algebra the following results are obtained:

$$V = \frac{1}{\alpha^2}\left((2m\zeta)^{\frac{1}{2}}\beta\zeta^{\frac{1}{2}} + 4\left(\frac{T}{\Theta}\right)^5 \mathscr{J}_5 \frac{m^3 D}{\hbar^3\Lambda}\right), \tag{46}$$

$$X = \frac{1}{\alpha^2}\left[\tfrac{1}{3}\pi^2(2m\zeta)^{\frac{1}{2}}\beta\zeta^{\frac{1}{2}} + 2\left(\frac{T}{\Theta}\right)^3\left\{\mathscr{J}_5 + \frac{D}{\zeta}\left(\frac{T}{\Theta}\right)^2\left(\frac{2\pi^2}{3}\mathscr{J}_5 - \tfrac{1}{3}\mathscr{J}_7\right)\right\}\frac{m^3\zeta}{\hbar^3\Lambda}\right], \tag{47}$$

$$W = \frac{1}{\alpha}(2m\zeta)^{\frac{1}{2}}, \quad Y = \frac{\pi^2}{3\alpha}(2m\zeta)^{\frac{1}{2}}. \tag{48}$$

These expressions are exactly the leading terms in (42) to (45) if they are expanded in powers of $1/\alpha^2$, but, whereas equations (42) to (45) are only valid at low temperatures, equations (46) to (48) are valid at all temperatures. This is a very unexpected result,

since it is known that the higher order correction terms are important in equations (42) to (45) and that the effects of the lattice vibrations and of the impurities are not strictly additive This matter is discussed in more detail in § 5.

4. The electrical conductivity

4·1. At high temperatures a time of relaxation exists, and the electrical conductivity can readily be calculated by means of equations (19) and (27). The resulting expression for σ can be put into various forms with a certain amount of algebraic manipulation and by using the relations (Wilson 1936, p 16, equation (29))

$$n_s = \frac{(2m_s\zeta)^{\frac{1}{2}}}{3\pi^2\hbar^3}, \quad n_d = \frac{(2m_d\zeta')^{\frac{1}{2}}}{3\pi^2\hbar^3}, \tag{49}$$

where n_s and n_d are the number of electrons and vacancies in the s- and d-bands respectively. The most convenient form for σ seems to be

$$\sigma = \epsilon^2 \frac{\left(\frac{n_s\tau_s}{m_s} + \frac{n_d\tau_d}{m_d}\right)^2 + \left(\frac{\epsilon H}{c}\right)^2 (n_s - n_d)^2 \frac{\tau_s^2\tau_d^2}{m_s^2 m_d^2}}{\frac{n_s\tau_s}{m_s} + \frac{n_d\tau_d}{m_d} + \left(\frac{\epsilon H}{c}\right)^2 \left(\frac{n_s\tau_d}{m_d} + \frac{n_d\tau_s}{m_s}\right) \frac{\tau_s\tau_d}{m_s m_d}} \tag{50}$$

In the absence of a magnetic field the conductivity is

$$\sigma_0 = \epsilon^2 \left(\frac{n_s\tau_s}{m_s} + \frac{n_d\tau_d}{m_d}\right), \tag{51}$$

and by combining (50) and (51) it is seen that the relative change in resistivity is

$$\frac{\rho - \rho_0}{\rho_0} = \frac{\sigma_0 - \sigma}{\sigma} = \frac{\left(\frac{\epsilon H}{c}\right)^2 \frac{n_s\tau_s}{m_s} \frac{n_d\tau_d}{m_d} \left(\frac{\tau_s}{m_s} + \frac{\tau_d}{m_d}\right)^2 \left(\frac{n_s\tau_s}{m_s} + \frac{n_d\tau_d}{m_d}\right)^{-2}}{1 + \left(\frac{\epsilon H}{c}\right)^2 (n_s - n_d)^2 \frac{\tau_s^2}{m_s^2} \frac{\tau_d^2}{m_d^2} \left(\frac{n_s\tau_s}{m_s} + \frac{n_d\tau_d}{m_d}\right)^{-2}} \tag{52}$$

The question now arises whether a more general formula of the same type can be found for the change in resistance which will be valid at all temperatures From previous work on the electrical resistance in the absence of a magnetic field (Wilson 1937, Dube 1938), it is known that the answer is no if an attempt is made to find an exact formula, but it is still possible to find an 'interpolation formula' which has the same validity as the well-known 'interpolation formula' for the normal electrical resistance. The obvious way to set up such an interpolation formula is to eliminate τ_s and τ_d from equations (50) to (52) and the equations

$$\sigma_{0s} = n_s\epsilon^2\tau_s/m_s, \quad \sigma_{0d} = n_d\epsilon^2\tau_d/m_d, \tag{53}$$

all of which are only valid at high temperatures Since the times of relaxation do not appear explicitly in the resulting expressions, but are replaced by the conductivities in zero magnetic field, which of course always have a direct physical significance, it seems likely that the formulae so obtained will have a wider validity than their derivation would suggest. This conjecture turns out to be well founded,

since the interpolation formulae reduce to the correct formulae in the three limiting cases of high temperature and arbitrary magnetic field, low temperature and arbitrary magnetic field, and large magnetic field and arbitrary temperature It would be difficult to set up a simple interpolation formula which satisfied further criteria than reducing to the correct expressions in the above three limiting cases

4·2 The simplest way to prove the statements made in the preceding paragraph is to compare the various expressions for V and W. It can readily be seen that, replacing $m/(ne^2\tau)$ in the expression (27) for V by

$$\frac{1}{\sigma_0} = \rho_0 = \rho_r + \rho_i = \frac{3\pi^2\hbar^4\beta}{(2m)^{\frac{1}{2}}\epsilon^2\zeta} + \frac{3\pi^2\hbar D}{2\epsilon^2\zeta^3\Lambda}\left(\frac{T}{\Theta}\right)^5 \mathscr{J}_5 \tag{54}$$

and treating all terms involving $\mathscr{J}_5$ as small compared with the others, the expression (42) is obtained as the first approximation Now (54) is the usual interpolation formula for the resistivity of a metal containing a single band, assuming Matthiessen's rule concerning the additivity of the residual and ideal resistivities ρ_r and ρ_i (see Wilson 1937, p. 378 and § 5 *infra*) Hence the expression

$$V = \frac{(2m\zeta)^{\frac{1}{2}}}{\hbar}\frac{ne^2\rho_0}{\alpha^2 + n^2\epsilon^4\rho_0^2/\hbar^2} \tag{55}$$

reduces at high temperatures to (27) and at low temperatures to (42) Also, for very high magnetic fields, (55) reduces to

$$V = \frac{(2m\zeta)^{\frac{1}{2}}}{\hbar}\frac{ne^2}{\alpha^2}\rho_0,$$

which is exactly (46), assuming that ρ_0 is given by (54) at all temperatures. In the general case the expression (55) is to be used for V and the corresponding expression

$$W = (2m\zeta)^{\frac{1}{2}}\frac{\alpha}{\alpha^2 + n^2\epsilon^4\rho_0^2/\hbar^2} \tag{56}$$

for W, with ρ_0 given by (54), all quantities having a suffix s or d when required.

4 3 The generalized expressions which replace (50), (51) and (52) are

$$\sigma = \frac{(\sigma_{0s}+\sigma_{0d})^2 + \left(\frac{H}{\epsilon c}\right)^2\frac{(n_s-n_d)^2}{n_s^2 n_d^2}\sigma_{0s}^2\sigma_{0d}^2}{\sigma_{0s}+\sigma_{0d} + \left(\frac{H}{\epsilon c}\right)^2\left(\frac{\sigma_{0s}}{n_s^2}+\frac{\sigma_{0d}}{n_d^2}\right)\sigma_{0s}\sigma_{0d}}, \tag{57}$$

$$\sigma_0 = \sigma_{0s} + \sigma_{0d}, \tag{58}$$

$$\frac{\sigma_0 - \sigma}{\sigma} = \frac{\left(\frac{H}{\epsilon c}\right)^2\left(\frac{\sigma_{0s}}{n_s}+\frac{\sigma_{0d}}{n_d}\right)^2\frac{\sigma_{0s}\sigma_{0d}}{(\sigma_{0s}+\sigma_{0d})^2}}{1 + \left(\frac{H}{\epsilon c}\right)^2\frac{(n_s-n_d)^2}{n_s^2 n_d^2}\frac{\sigma_{0s}^2\sigma_{0d}^2}{(\sigma_{0s}+\sigma_{0d})^2}}, \tag{59}$$

σ_{0s} and σ_{0d} being given by expressions like (54) but containing the appropriate suffixes s and d.

These equations have been given before but only in the form (50) to (52), and Jones (1936) has used them to discuss the magnetic effects in bismuth Until now, however, the derivations given have all assumed the existence of a time of relaxation, and the equations could not be used with any confidence at low temperatures where the magnetic effects are most interesting and most marked There are two cases to discuss depending on whether n_s is or is not equal to n_d

Case (i) $n_s = n_d = \frac{1}{2}n$

In this case $(\rho-\rho_0)/\rho_0$ is proportional to H^2 for all field strengths Kapitza's experiments, on the other hand, indicate that while $(\rho-\rho_0)/\rho_0$ is proportional to H^2 for small fields it tends to increase more slowly than H^2 for large fields and in fact approximately linearly This indicates that case (i) is not realized in practice

In order to obtain simple qualitative results from the formulae it is sufficient to put $\rho_{0s} = \rho_{0d} = 2\rho_0$. Then

$$\rho = \rho_0 + \left(\frac{H}{nec}\right)^2 \frac{1}{\rho_0}. \tag{60}$$

If ρ is treated as a function of ρ_0, it has a minimum value $\rho_{\min} = 2H/(nec)$ which occurs when $\rho_0 = H/(nec)$. Now ρ_0 can never be less than the residual resistivity ρ_r, so that the minimum only occurs provided that $H > nec\rho_r$. Therefore the qualitative picture of ρ as a function of T is as follows

For $H < nec\rho_r$ the resistivity increases steadily with T from its value

$$\rho_r + \left(\frac{H}{nec}\right)^2 \frac{1}{\rho_r}$$

at $T = 0$, and for sufficiently high temperatures it differs negligibly from the ideal resistance ρ_i with the magnetic fields that are obtainable in practice

For $H > nec\rho_r$ the resistivity at first decreases as T increases (the decrease being initially $\rho_i\{(H/nec\rho_r)^2 - 1\}$ and hence proportional to T^5), reaches the minimum value $2H/(nec)$ at that value of T for which $\rho_0 = H/(nec)$ and then steadily increases

Case (ii) $n_s \neq n_d$

In this case $(\rho-\rho_0)/\rho_0$ is proportional to H^2 for small fields but reaches a constant value for very large fields In fields of intermediate value $(\rho-\rho_0)/\rho_0$ increases more slowly than H^2, and it is not difficult to reproduce the type of result obtained by Kapitza provided it is assumed that the fields which produce saturation lie beyond the experimental range.

Writing, for simplicity, $\rho_{0s} = \rho_{0d} = 2\rho_0$, (57) becomes

$$\rho = \rho_0 \frac{\rho_0^2 + \frac{1}{8}\left(\frac{H}{ec}\right)^2\left(\frac{1}{n_s^2} + \frac{1}{n_d^2}\right)}{\rho_0^2 + \frac{1}{16}\left(\frac{H}{ec}\right)^2 \frac{(n_s - n_d)^2}{n_s^2 n_d^2}}, \tag{61}$$

from which the temperature variation of ρ can be determined. For sufficiently small values of $n_s - n_d$, the temperature variation of ρ is the same as in case (i), namely,

a minimum exists for large values of H. If, however, $n_s - n_d$ is sufficiently large, there is no minimum, whatever the value of H, and ρ increases steadily with T.

The minimum in the $\rho - T$ curve for large values of H has been observed by Milner (1937) for cadmium There are also indications that other workers would have found similar results if they had pushed their experiments far enough, but Milner's work is the only one known to us in which the minimum has definitely been found. Milner gives no explanation of the existence of the minimum, but points out that it may be connected with the quantization of the electronic orbits due to the magnetic field It is certainly true that the conditions under which the Boltzmann equation are valid begin to break down when H is of the order of $nec\rho$ and that the quantization must then be considered, but without detailed calculation it is impossible to tell how large a correction to the present formulae is introduced by considering the quantization. This point will be considered in detail in a later paper, but for the present it need only be remarked that the existence of the minimum is shown by the elementary theory in which quantization is neglected, and it is probable that the more refined theory will only change the quantitative and not the qualitative results

Any detailed comparison of the present formulae with the experimental results is out of the question, first because of the extremely meagre and unsatisfactory nature of the experimental data available, and secondly because it would be unreasonable to expect the simplified model used to be exactly applicable to any real metal. The general form of the equations is, however, probably reasonably correct (for a discussion of the results to be expected from a general model see Peierls 1931), though no reliance can be placed upon the values of the numbers which occur in the formulae To explain most properties of metals it is sufficient to treat the electrons as if their energy levels were given by the simple formula $E = \hbar^2 |\mathbf{k}|^2/2m$, but if this is done there is no magneto-resistance effect. This is because the simpler phenomena depend essentially upon the total number of electrons present, whereas the magneto-resistance effect depends upon an average of the departure of the energy levels from spherical symmetry. In addition to n, the total number of electrons and vacancies, a new quantity n_a should therefore be introduced, which has the dimensions of a number of electrons per unit volume and which is related to the non-spherical nature of the surfaces of constant energy In the present model, $n = n_s + n_d$, while n_a^2 must be related to $n_s n_d$, since it vanishes if either $n_s = 0$ or $n_d = 0$. It therefore seems probable on general grounds that if more general models could be worked out (and at the moment there seems to be no prospect of this being achieved) formulae would be obtained similar to those derived here, but with n_s and n_d replaced by appropriate combinations of n and n_a and with the numerical factors changed In trying to interpret the experimental results it is best to assume that n_s and n_d are parameters connected with n and n_a in some not too clearly defined way, and not to use the two-band model in too literal a manner

The formulae given here are in fair quantitative agreement with the experimental results on cadmium For example, for one of Milner's specimens $\rho_0 = 4\cdot5 \times 10^{-21}$ gaussian units at 2·35° K, and $\rho = 1\cdot4 \times 10^{-18}$ in a field of 2×10^4 gauss Equation (60)

gives $n = 1{\cdot}7 \times 10^{22}$, and with this value it is estimated that in a field of $2{\cdot}28 \times 10^4$ gauss the resistivity of cadmium reaches a minimum value of $1{\cdot}8 \times 10^{-19}$ at a temperature corresponding to $\rho_0 = 0{\cdot}9 \times 10^{-19}$. The observed values are $2{\cdot}6 \times 10^{-19}$ and $0{\cdot}6 \times 10^{-19}$ respectively

5 Matthiessen's rule

5·1. Matthiessen's empirical rule states that the electrical resistivity consists of two parts, the ideal resistivity ρ_i characteristic of the pure metal, and the residual resistivity ρ_r which is due to impurities, strains, etc , and which varies from specimen to specimen Since the total resistivity ρ is supposed to be the sum of ρ_i and ρ_r, and since ρ_r is independent of temperature, all the resistance-temperature curves for a given metal in different states of purity should be capable of being superposed by shifting them parallel to the ρ-axis

Until lately Matthiessen's rule was supposed to be well founded, but recent experimental and theoretical work (see, for example, Gruneisen 1933, Justi & Scheffers 1938; and Dube 1938) has shown that the rule can only be considered to be a good first approximation, and that deviations from it are to be expected in those temperature regions where the residual and ideal resistances are of the same order of magnitude For metals of high purity this means temperatures of the order of 10° K

Dube (1938) attempted to determine the deviations from Matthiessen's rule for the simplest model in which the electrons are free and are located in a single band of normal form. He was able to obtain the first correction term, but owing to the complexity of the higher order corrections his formulae were not valid in the interesting temperature region His results, however, are of importance in showing that even with the simplest model deviations from Matthiessen's rule are to be expected

So far little attention has been given to the validity or otherwise of Matthiessen's rule when a magnetic field is present, but, unless care is taken to separate the various effects, much of the experimental work will be of little value. For the reasons given below Matthiessen's rule cannot be expected to be valid except more or less by accident, and attempts to simplify the experimental results by assuming the validity of the rule are not likely to be successful

For the two-band model used here Matthiessen's rule does not hold even in the absence of a magnetic field Since the electrons in the two bands are independent of one another the conductivities of the two bands are additive and

$$\sigma = \sigma_s + \sigma_d \tag{62}$$

Neglecting the effect discussed by Dube and assuming that Matthiessen's rule applies to each band separately,

$$\rho_s = \rho_{rs} + \rho_{is}, \quad \rho_d = \rho_{rd} + \rho_{id} \tag{63}$$

It is not then possible to express ρ as the sum of ρ_r and ρ_i where ρ_r depends only on ρ_{rs} and ρ_{rd}, while ρ_i depends only on ρ_{is} and ρ_{id}

To find the deviations from Matthiessen's rule the procedure is as follows Put

$$\rho_{rs} = a, \quad \rho_{rd} = b, \quad \rho_{is} = x, \quad \rho_{id} = y. \tag{64}$$

Then

$$\rho_s = a + x, \quad \rho_d = b + y \tag{65}$$

and

$$\rho = \frac{(a+x)(b+y)}{a+b+x+y}. \tag{66}$$

Now define ρ_r as the resistivity when $x = y = 0$, and ρ_i as the resistivity when $a = b = 0$, that is,

$$\rho_r = \frac{ab}{a+b}, \quad \rho_i = \frac{xy}{x+y}. \tag{67}$$

Then the deviation Δ from Matthiessen's rule is

$$\Delta = \rho - (\rho_r + \rho_i) = \frac{(bx - ay)^2}{(a+b)(x+y)(a+b+x+y)} \tag{68}$$

Having found ρ_i by measurements on the ideally pure substance, the apparent residual resistivity of any specimen is $\rho_r' = \rho - \rho_i = \rho_r + \Delta$ Equation (68) shows that Δ is always positive and is strongly temperature dependent

To obtain a qualitative picture of the temperature variation of Δ it is sufficient to put $\lambda a = b$ and $\mu x = y$ Then

$$\Delta = \frac{ax(\lambda - \mu)^2}{(1+\lambda)(1+\mu)\{a(1+\lambda) + x(1+\mu)\}}.$$

At low temperatures x is small compared with a, so that Δ is proportional to x, that is, to ρ_i, while at high temperatures, where x is large compared with a, Δ is proportional to a, that is, to ρ_r. It is therefore seen that Δ is zero when $T = 0$, it increases as T increases and finally reaches a constant value proportional to ρ_r For very pure metals the constancy of Δ holds down to very low temperatures, and the constancy of the apparent residual resistance is no guarantee of the validity of Matthiessen's rule unless the measurements are carried out at temperatures where ρ_i is of the same order of magnitude as ρ_r.

The observed values of Δ are positive, increase with temperature and are smallest for the purest specimens (Grüneisen 1933) This is in agreement with the predictions of the two-band model and also probably of any model more complicated than the one-band free-electron model The Δ calculated by Dube for the latter model is negative and proportional to $1/\rho_r$. Its importance is presumably small compared with that of the effects discussed here

5·2 Since Matthiessen's rule does not hold in the absence of a magnetic field it clearly cannot hold when a magnetic field is present The results are even more complicated than when the magnetic field is zero, and no useful purpose would be served by discussing them in detail except in the limiting case of an infinite magnetic field, since the results are then very much simplified Further, when H is infinite

the equations derived in § 3·3 are exact, so that the results given here are valid for all temperatures without approximation. As in § 4·3 there are two cases to consider.

Case (i). $n_s = n_d$

In this case equations (46) and (48) lead to

$$\sigma = \frac{e^2}{3\pi^2\hbar^4\alpha^2}\left[(2m_s\zeta)^{\frac{1}{2}}\beta_s\zeta^{\frac{1}{2}} + (2m_d\zeta')^{\frac{1}{2}}\beta_d\zeta'^{\frac{1}{2}} + 4\left(\frac{T}{\Theta}\right)^5 \mathscr{J}_5\left(\frac{m_s^3 D_s}{\hbar^3 \Lambda_s} + \frac{m_d^3 D_d}{\hbar^3 \Lambda_d}\right)\right], \quad (69)$$

which is a remarkable result since it shows that the residual and ideal conductivities are additive and not the resistivities. This represents as extreme a departure from Matthiessen's rule as it is possible to imagine. It is a consequence of the fact that, when $n_s = n_d$, $\sigma_{H=\infty}$ is proportional to $\rho_{H=0}$.

Case (ii). $n_s \neq n_d$

In this case equations (46) and (48) give

$$\frac{1}{\sigma} = \frac{3\pi^2\hbar^4}{e^2}\left[(2m_s\zeta_s)^{\frac{1}{2}}\beta_s\zeta^{\frac{1}{2}} + (2m_d\zeta')^{\frac{1}{2}}\beta_d\zeta'^{\frac{1}{2}} + 4\left(\frac{T}{\Theta}\right)^5 \mathscr{J}_5\left(\frac{m_s^3 D_s}{\hbar^3 \Lambda_s} + \frac{m_d^3 D_d}{\hbar^3 \Lambda_d}\right)\right] \times \{(2m_s\zeta)^{\frac{1}{2}} - (2m_d\zeta')^{\frac{1}{2}}\}^{-1}, \quad (70)$$

and Matthiessen's rule holds exactly.

Since these two particular cases give such widely differing results, it is clear that nothing simple can be expected in the general case.

6. The electronic thermal conductivity

6·1. When a time of relaxation exists, equations (28) give $X = \frac{1}{3}\pi^2 V$, $Y = \frac{1}{3}\pi^2 W$. Hence the Wiedemann-Franz law holds for all values of the magnetic field, and all results concerning the electrical conductivity can be translated at once into corresponding results concerning the electronic thermal conductivity. To obtain results of more general validity the procedure as in § 4 is to set up an interpolation formula which reduces to the correct form in the three limiting cases for which the solution is known. Here, however, the situation is more delicate than for the electrical conductivity, since it is known that the extrapolation of results based upon the existence of a time of relaxation gives an incorrect answer at low temperatures, but with the three limiting cases as guide it is not difficult to find the correct generalization. It is merely necessary to replace everywhere ρ_0 by $L_n T/\kappa_0$, where κ_0 is the thermal conductivity in zero magnetic field and L_n is the 'normal' value $\frac{1}{3}(\pi k/e)^2$ of the Lorenz number The expressions corresponding to (55) and (56) are therefore

$$X = \frac{\pi^2(2m\zeta)^{\frac{1}{2}}}{3\hbar} \frac{ne^2 L_n T/\kappa_0}{\alpha^2 + n^2e^4 L_n^2 T^2/\hbar^2\kappa_0^2}, \quad (71)$$

$$Y = \frac{1}{3}\pi^2(2m\zeta)^{\frac{1}{2}} \frac{\alpha}{\alpha^2 + n^2e^4 L_n^2 T^2/\hbar^2\kappa_0^2}, \quad (72)$$

the appropriate suffixes s and d being understood. It is further assumed that Matthiessen's rule holds for the thermal resistivity and hence (Wilson 1937, p. 378) that for each band

$$\frac{1}{\kappa_0}=\frac{1}{\kappa_r}+\frac{1}{\kappa_i}=\frac{9\hbar^4\beta}{(2m)^{\frac{1}{2}}\zeta k^{\frac{1}{2}}T}+\frac{27\hbar}{4\pi^2\zeta^2k^2\Lambda\Theta}\left(\frac{T}{\Theta}\right)^2\left\{\mathscr{J}_5+\frac{D}{\zeta}\left(\frac{T}{\Theta}\right)^2\left(\frac{2\pi^2}{3}\mathscr{J}_3-\tfrac{1}{3}\mathscr{J}_7\right)\right\}. \quad (73)$$

To show that (71) and (72) give the correct expressions for X and Y it is merely necessary to insert the value for κ_0 given by (73) and show that (44), (45), (47) and (48) are obtained in the two limiting cases of $1/\kappa_i$ very small and H very large. The proof is elementary and the details are left to the reader.

It is now possible to write down the expressions corresponding to (57), (58) and (59) They are

$$\kappa=\frac{(\kappa_{0s}+\kappa_{0d})^2+\left(\frac{3H\epsilon}{\pi^2k^2cT}\right)^2\frac{(n_s-n_d)^2}{n_s^2n_d^2}\kappa_{0s}^2\kappa_{0d}^2}{\kappa_{0s}+\kappa_{0d}+\left(\frac{3H\epsilon}{\pi^2k^2cT}\right)^2\left(\frac{\kappa_{0s}}{n_s^2}+\frac{\kappa_{0d}}{n_d^2}\right)\kappa_{0s}\kappa_{0d}}, \quad (74)$$

$$\kappa_0=\kappa_{0s}+\kappa_{0d}, \quad (75)$$

$$\frac{\kappa_0-\kappa}{\kappa}=\frac{\left(\frac{3H\epsilon}{\pi^2k^2cT}\right)^2\left(\frac{\kappa_{0s}}{n_s}+\frac{\kappa_{0d}}{n_d}\right)^2\frac{\kappa_{0s}\kappa_{0d}}{(\kappa_{0s}+\kappa_{0d})^2}}{1+\left(\frac{3H\epsilon}{\pi^2k^2cT}\right)^2\frac{(n_s-n_d)^2}{n_s^2n_d^2}\frac{\kappa_{0s}^2\kappa_{0d}^2}{(\kappa_{0s}+\kappa_{0d})^2}}. \quad (76)$$

6·2 The experimental results on the thermal conductivity are usually expressed in terms of the Lorenz number $L=\kappa/\sigma T$ The experimental and theoretical values for the Lorenz number in the absence of a magnetic field have been discussed in detail by Makinson (1938), and it is now necessary to consider how his results have to be supplemented to take account of the effect of a magnetic field

The Lorenz number is given at once by equations (57) and (74), and the results can easily be worked out for any particular case. In order to illustrate the general characteristics of L, various simplifying assumptions are introduced, namely, $\sigma_{0s}=\sigma_{0d}=\frac{1}{2}\sigma_0$, $\kappa_{0s}=\kappa_{0d}=\frac{1}{2}\kappa_0$, and the cases $n_s=n_d$ and $n_s\neq n_d$ are considered separately

The value of L when $H=0$ is denoted by L_0, and the 'normal' value of L by L_n, which is the value of L_0 when a time of relaxation exists. The general characteristics of $L_0=\kappa_0/\sigma_0T$ are shown in figure 3 of Makinson's paper. (Note that his notation differs slightly from that given here.) Disregarding metals in which the number of free electrons is very small (which includes bismuth), L_0 is equal to L_n at both high and low temperatures but may be considerably less at intermediate temperatures.

Case (i) $n_s=n_d=\frac{1}{2}n$

In this case

$$L=\frac{\kappa_0}{\sigma_0T}\frac{1+(H/nec)^2\sigma_0^2}{1+(3H\epsilon/\pi^2k^2ncT)^2\kappa_0^2}. \quad (77)$$

For large values of H this reduces to

$$L_{H=\infty} = \left(\frac{\pi^2 k^2}{3e^2}\right)^2 \frac{\sigma_0 T}{\kappa_0} = \frac{L_n^2}{L_0}. \tag{78}$$

Since in general $L_0 \leqslant L_n$, it is seen from (77) and (78) that, as H increases from zero to infinity, L increases steadily from the value L_0, which is less than L_n, to the value L_n^2/L_0, which is greater than L_n. If, on the other hand, $L_0 \geqslant L_n$, L decreases as H increases.

Case (II). $n_s \neq n_d$.

In this case

$$L = \frac{\kappa_0}{\sigma_0 T} \frac{1 + \frac{1}{8}\left(\frac{H}{ec}\right)^2 \left(\frac{1}{n_s^2} + \frac{1}{n_d^2}\right) \sigma_0^2}{1 + \frac{1}{8}\left(\frac{3He}{\pi^2 k^2 cT}\right)^2 \left(\frac{1}{n_s^2} + \frac{1}{n_d^2}\right) \kappa_0^2} \frac{1 + \frac{1}{16}\left(\frac{3He}{\pi^2 k^2 cT}\right)^2 \frac{(n_s - n_d)^2}{n_s^2 n_d^2} \kappa_0^2}{1 + \frac{1}{16}\left(\frac{H}{ec}\right)^2 \frac{(n_s - n_d)^2}{n_s^2 n_d^2} \sigma_0^2}. \tag{79}$$

For small values of H this reduces to

$$L = L_0 \left\{1 + \frac{1}{16}\left(\frac{H}{ec}\right)^2 \frac{(n_s + n_d)^2}{n_s^2 n_d^2} \left(1 - \frac{L_0^2}{L_n^2}\right) \sigma_0^2\right\}, \tag{80}$$

which shows that L increases initially with H if $L_0 < L_n$. For very large values of H, however, L has the value L_0, so that L initially increases with H, reaches a maximum and finally tends asymptotically to its initial value L_0.

The experimental data with which to compare the theoretical results are not very extensive (Gruneisen & Adenstedt 1938, Gruneisen & Erfling 1940), but they are in qualitative agreement with the predictions outlined above. The Lorenz number is increased by the application of a magnetic field, but there is no evidence as to whether it passes through a maximum or not. Measurements on metals which do and do not show saturation in strong fields are particularly desirable to settle this point.

To verify that the formulae give results of the right order of magnitude, equations (60) and (77) have been used to obtain two estimates of n for beryllium. A typical set of observed values is (Gruneisen & Erfling 1940) $\rho_0 = 8.4 \times 10^{-21}$, $L_0 = 1{\cdot}42 \times 10^{-13}$ at 22.6° K and $\rho = 7.9 \times 10^{-19}$, $L = 2{\cdot}92 \times 10^{-13}$ at the same temperature in a field of $1{\cdot}17 \times 10^4$ gauss. (All figures are in gaussian units.) With these values equation (60) gives $n = 10^{23}$ and equation (77) gives $n = 6{\cdot}2 \times 10^{22}$. Better agreement could of course have been obtained by using equations (61) and (79) which have two adjustable constants.

7. The lattice conductivity

In poor conductors such as bismuth an appreciable part of the thermal conductivity is due to the crystal lattice. This matter has been discussed at length by Makinson (1938), and it is merely necessary to complete his discussion by considering

whether the lattice conductivity is affected by a magnetic field or not There is, of course, no direct effect, but in certain temperature regions the lattice conductivity is almost entirely governed by the collisions of the electrons with the lattice, and as the paths of the electrons are changed by the magnetic field there might be an indirect effect on the lattice conductivity It turns out, however, that there is no effect

To prove this it is merely necessary to re-examine Makinson's argument that the interaction of the lattice with the electrons depends solely upon the equilibrium constants of the lattice provided that no electric current is flowing. Makinson's argument, given below his equation (18), depends only on the fact that in the absence of an electric current $c(\eta)$ is an odd function of η ($c(\eta)$ determines the departure of the distribution function from the Fermi function). This is true whether there is a magnetic field or not, the one extra complication introduced by the presence of a magnetic field being that there are two functions $c_1(\eta)$ and $c_2(\eta)$ to consider instead of one, but both are still odd functions of η when the electric current is zero. A magnetic field is therefore without effect upon the conductivity of the lattice.

This result justifies to a certain extent one of the principal methods of separating the lattice conductivity from the total thermal conductivity. Both κ and σ are measured in a magnetic field of increasing intensity at a given temperature, and κ is plotted against σT (Gruneisen & Adenstedt 1938) The resulting curve is extrapolated to zero value of σT, and the intercept on the κ axis gives the lattice conductivity. There are two assumptions here first, that the lattice conductivity is independent of the magnetic field, and secondly, that the electrical conductivity and the electronic thermal conductivity are both zero when the magnetic field is infinite. It has just been seen that the first assumption is correct, but it is by no means certain that the second is justified. For the model discussed here both σ and κ show saturation in strong fields (unless $n_s = n_d$), and it is not then correct to extrapolate to $\sigma = 0$ Instead the curve should be extrapolated to its end-point at $\sigma = \sigma_{H=\infty}$, which would give a somewhat larger value for the lattice conductivity than extrapolating to $\sigma = 0$ This complication does not seem to have been taken into account in discussing the experimental results, it would only affect them if high accuracy is required *

References

Bronstein, M 1932 *Phys. Z. Sowjet.* **2**, 28.
Dube, G P. 1938 *Proc Camb Phil Soc* **34**, 559
Grunеisen, E 1933 *Ann Phys , Lpz ,* (5), **16**, 530
Grüneisen, E & Adenstedt, H. 1938 *Ann Phys , Lpz.,* (5), **31**, 714
Grunеisen, E & Erfling, H. D 1940 *Ann. Phys , Lpz ,* (5), **38**, 399.
Harding, J. W. 1933 *Proc. Roy. Soc.* A, **140**, 205

* Extrapolation of the $\kappa - \sigma T$ curve to $\sigma_{H=\infty}$ gives $\kappa_g + \kappa_{H=\infty}$, where κ_g is the lattice thermal conductivity and κ_H is the electronic thermal conductivity when the magnetic field is H, but equation (79) shows that $L = L_0$ for both $H = 0$ and $H = \infty$ when $n_+ \neq n_-$. Hence $\kappa_g = \kappa - \kappa_H = \kappa - L\sigma_H T$ is the intercept on the κ-axis of the straight line joining the end-points of the $\kappa - \sigma T$ curve at $H = 0$ and $H = \infty$.

Jones, H. 1936 *Proc Roy. Soc.* A, **155**, 653.
Justi, E. & Scheffers, H. 1938 *Phys Z.* **39**, 591.
Kohler, M. 1940 *Ann Phys., Lpz*, (5), **38**, 283
Kohler, M. 1941 *Ann Phys, Lpz.*, (5), **40**, 601.
Makinson, R. E. B 1938 *Proc. Camb Phil. Soc* **34**, 474.
Milner, C. J 1937 *Proc. Roy. Soc.* A, **160**, 207.
Peierls, R. 1931 *Ann. Phys., Lpz*, (5), **10**, 97.
Titeica, S 1935 *Ann. Phys., Lpz.*, (5), **22**, 129.
Wilson, A. H 1936 *The theory of metals.* Cambridge.
Wilson, A H 1937 *Proc. Camb. Phil. Soc* **33**, 371.
Wilson, A. H 1938 *Proc. Roy Soc.* A, **167**, 580.

A general kinetic theory of liquids
III. Dynamical properties

By M. Born, F.R S. and H. S. Green

(*Received* 4 *December* 1946—*Read* 26 *June* 1947)

The theory of liquids formulated in part I and applied to the equilibrium state in part II is here extended to liquids in motion The connexion between the macroscopic and microscopic properties is revealed by the derivation of a set of generalized hydrodynamical equations, of which the fundamental equations of hydrodynamics are a special case; the more general equations describe the mean motion of clusters of molecules in the fluid It is shown that the pressure tensor and energy-flux vector in a fluid consist of two parts, due to the thermal motion of the molecules and the intermolecular forces respectively, of which only the first is found in the kinetic theory of gases, but of which the second is dominant for the liquid state

A method is evolved for the study of those 'normal' non-uniform states which relate to actual monatomic fluids in motion It becomes apparent, as in the case of equilibrium, that there is a region of temperature and density where analytical singularities arise, closely associated with the process of condensation.

Rigorous expressions for the coefficients of viscosity and thermal conduction are then derived which apply equally to the liquid and the gas They consist of two parts due to the thermal motion and molecular forces respectively, of which the first is dominant for the gas, and the second for the liquid By approximating to the rigorous formula, an expression for the viscosity of liquids is obtained, comparable with certain other formulae, previously proposed on quasi-empirical grounds, and giving good agreement with experiment

An integro-differential equation is derived for the determination of the distribution functions relating to the non-uniform state A full discussion is given of the simplest case, and the velocity distribution in non-uniform liquids and gases examined

1. Introduction

In the first paper of this series (Born & Green 1946), we have explained a general method for dealing with an assembly of equal particles not in equilibrium, in the second (Green 1947), the special case of equilibrium has been treated, and it has been shown that even here, where the new method is completely equivalent to statistical mechanics, it is more powerful in obtaining explicit results. In the present paper, a moving fluid will be treated in more detail, and not only the thermo-mechanical

equations of motion, but also rigorous expressions for the coefficients of viscosity and thermal conduction will be given. The latter will be simplified for comparison with expressions found in the literature (which, however, can be regarded only as dimensional formulae) In this way theoretical justification for certain empirical formulae will be found, of the type first mentioned by de Guzman (1913) and studied in detail by Andrade (1934), which are known to give excellent agreement with experiment.

Concerning the method of approximation, it has been found that the method developed by Hilbert (1912), Enskog (1911, 1917) and Chapman (1912, 1916), fails for the multiple distribution functions ($h > 1$). A new one has therefore been devised, based on an expansion with respect to the space gradients of the equilibrium parameters (density, mean velocity and temperature)

One feature of the results has to be mentioned particularly, namely, the appearance of condensation as a singularity not only in the equilibrium equations, but also in the expressions for the dynamical constants. We wish to stress the point that we differ fundamentally from the ideas about condensation held by many modern writers, and clearly formulated by Kahn (1938) in his well-known dissertation, namely, that the horizontal part of the isotherm in the p-v diagram (p = const), representing the co-existence of liquid and gas, is contained in the general thermodynamical expressions. In fact, the co-existence of the two phases means only that there are two values of v for the same p, the straight line connecting these points has no meaning unless a parameter strange to the theory of one phase is introduced, e g the mass ratio of the two phases. We believe that Kahn's derivation of this line p = const is based on a misconception of the situation; for he assumes the virial coefficients (β_n) to be constants, whereas they depend on the volume for high values of n. Kahn's results have really a different meaning he has found the liquid branch of the isotherm without noticing it.

In this paper no attempt will be made to obtain exact numerical values for physical constants, as this will require extended and tedious calculations. Apart from these, we hope to generalize the theory in the following different directions. The first and most obvious step is the introduction of internal degrees of freedom of the particles (molecules), this is probably closely connected with giving up the assumption of central forces. Then mixtures of two or more kinds of particles have to be considered in order to obtain the laws of ordinary and thermal diffusion, the latter (Soret effect) will be of special interest, as no satisfactory theory for liquids seems to exist Also our theory gives, as shown in part II, clear indications of the appearance of the solid state; we hope that it will lead to a better understanding of the present lattice dynamics and its limits, in particular to a theory of melting and plasticity.

The most important generalization envisaged is the introduction of quantum mechanics. For there are known several examples of quantum 'liquids': helium II (superfluidity), the substance of nuclei (droplet model), and perhaps electrons in metals (superconductivity). It is, of course, hopeless to try to understand the

abnormal viscosity in liquid helium if one has no theory of viscosity in normal liquids. This gap has now been closed, and we wish to extend our theory of viscosity to quantum liquids.

The tensor notation used in this paper is that devised by Chapman & Milne (see Chapman & Cowling 1939, chapter 1). Vectors are represented in clarendon type (**a**), and cartesian tensors in sans serif type (**w**). The transpose of a tensor **w** is $\overline{\mathbf{w}}$; its products with a vector **a** (both vectors) are denoted by **w**.**a** and **a**.**w**; the spur product with a second tensor **w**′ (a scalar), by **w** **w**′, etc.

2 Hydrodynamics

The fundamental equation of the theory [(4·5) of part I] may be written in the form

$$\frac{\partial f_h}{\partial t}+\sum_{i=1}^{h}\left(\boldsymbol{\xi}^{(i)}\cdot\frac{\partial f_h}{\partial \mathbf{x}^{(i)}}+\boldsymbol{\eta}_h^{(i)}\cdot\frac{\partial f_h}{\partial \boldsymbol{\xi}^{(i)}}\right)$$
$$=\frac{1}{m}\sum_{i=1}^{h}\int\int\left(\frac{\partial \phi^{(ih+1)}}{\partial \mathbf{x}^{(i)}}\cdot\frac{\partial}{\partial \boldsymbol{\xi}^{(i)}}\right)\left(f_{h+1}-\frac{f_h f_1^{(h+1)} n_{h+1}}{n_h n_1^{(h+1)}}\right)d\mathbf{x}^{(h+1)}\,d\boldsymbol{\xi}^{(h+1)}, \quad (2\cdot1)$$

where
$$m\boldsymbol{\eta}_h^{(i)}=-\sum_{j=1}^{h}\frac{\partial \phi^{(ij)}}{\partial \mathbf{x}^{(i)}}-\int\frac{\partial \phi^{(ih+1)}}{\partial \mathbf{x}^{(i)}}\frac{n_{h+1}}{n_h}d\mathbf{x}^{(h+1)}+\mathbf{P}^{(i)}, \quad (2\cdot2)$$

so that, while $\mathbf{P}^{(i)}=\mathbf{P}(t,\mathbf{x}^{(i)})$ is the *external* force acting on a molecule at $\mathbf{x}^{(i)}$ at time t, $m\boldsymbol{\eta}_h^{(i)}$ includes the average force due to the presence of the other molecules, the positions of the first h of which are assumed to be known.

In this section, the generalized hydrodynamical equations referring to such a cluster of h molecules will be derived from (2·1). By integrating this equation over the velocities $\boldsymbol{\xi}^{(1)}$.. $\boldsymbol{\xi}^{(h)}$,

$$\frac{\partial n_h}{\partial t}+\sum_{i=1}^{h}\frac{\partial}{\partial \mathbf{x}^{(i)}}\,(n_h\mathbf{u}_h^{(i)})=0 \quad (2\cdot3)$$

is obtained, where
$$n_h\mathbf{u}_h^{(i)}=\int_{\cdot}^{(h)}\int f_h\,\boldsymbol{\xi}^{(i)}\prod_{j=1}^{h}d\boldsymbol{\xi}^{(j)}. \quad (2\cdot4)$$

(2·3) and (2·4) were obtained previously, as (3·3) and (3·4) of part I, where (2·3) was interpreted as the generalized equation of continuity

Let I be any function of the time t and the co-ordinates $\mathbf{x}^{(1)}$ $\mathbf{x}^{(h)}$ Then it follows from (2·3) that

$$\frac{\partial}{\partial t}(n_h I)=n_h\frac{dI}{dt}-\sum_{i=1}^{h}\frac{\partial}{\partial \mathbf{x}^{(i)}}\cdot(n_h I\mathbf{u}_h^{(i)}), \quad (2\cdot5)$$

where $\dfrac{d}{dt}$ represents the operator $\dfrac{\partial}{\partial t}+\sum_{i=1}^{h}\mathbf{u}_h^{(i)}\cdot\dfrac{\partial}{\partial \mathbf{x}^{(i)}}$.

Next, after multiplying (2·1) by $\boldsymbol{\xi}^{(i)}$ and integrating over all the velocities,

$$\frac{\partial}{\partial t}(n_h\mathbf{u}_h^{(i)})+\sum_{j=1}^{h}\frac{\partial}{\partial \mathbf{x}^{(j)}}\cdot\left(\int^{(h)}\cdots\int f_h\,\boldsymbol{\xi}^{(j)}\boldsymbol{\xi}^{(i)}\prod_{k=1}^{h}d\boldsymbol{\xi}^{(k)}\right)=n_h\boldsymbol{\eta}_h^{(i)} \quad (2\cdot6)$$

is obtained, since the right-hand side of (2·1) is seen to contribute nothing, after an integration by parts over $\xi^{(i)}$. On applying (2·5), (2·6) becomes

$$mn_h \frac{d}{dt}\mathbf{u}_h^{(i)} + \sum_{j=1}^{h} \frac{\partial}{\partial \mathbf{x}^{(j)}} \cdot \mathsf{k}_h^{(ij)} = mn_h \eta_h^{(i)}, \tag{2·7}$$

where

$$\mathsf{k}_h^{(ij)} = m \int \cdots \int f_h \mathbf{v}_h^{(i)} \mathbf{v}_h^{(j)} \prod_{k=1}^{h} d\xi^{(k)}, \tag{2·8}$$

and

$$\mathbf{v}_h^{(i)} = \xi^{(i)} - \mathbf{u}_h^{(i)}. \tag{2·9}$$

Further, on writing

$$\mathsf{p}_h^{(ij)} = \mathsf{k}_h^{(ij)} + \mathsf{I}_h^{(ij)}, \tag{2·10}$$

where $\mathsf{I}_h^{(ij)}$ is defined by the equation

$$\sum_{j=1}^{h} \frac{\partial}{\partial \mathbf{x}^{(j)}} \cdot \mathsf{I}_h^{(ij)} = n_h \sum_{j=1}^{h} \frac{\partial \phi^{(ij)}}{\partial \mathbf{x}^{(i)}} + \int n_{h+1} \frac{\partial \phi^{(ih+1)}}{\partial \mathbf{x}^{(i)}} d\mathbf{x}^{(h+1)}, \tag{2·11}$$

(2·7) assumes the form

$$mn_h \frac{d}{dt}\mathbf{u}_h^{(i)} + \sum_{j=1}^{h} \frac{\partial}{\partial \mathbf{x}^{(j)}} \cdot \mathsf{p}_h^{(ij)} = n_h \mathbf{P}^{(i)}. \tag{2·12}$$

This is the generalization of the equation of motion of ordinary hydrodynamics, to which it reduces when $h = 1$. It is clear that the pressure tensor $\mathsf{p}_h^{(ij)}$ consists of two parts $\mathsf{k}_h^{(ij)}$ and $\mathsf{I}_h^{(ij)}$, of which the first is due to the thermal motion of the molecules, and the second to the intermolecular forces. It is customary to neglect this latter part in the kinetic theory of gases, but in liquids, far from being negligible, it is clearly the dominant part.

When (2·1) is multiplied by $\frac{1}{2}\xi^{(i)2}$ and integrated over all the velocities, the resulting equation is

$$\frac{\partial}{\partial t}\left(\frac{3k}{2m} n_h T_h^{(i)} + \tfrac{1}{2} n_h \mathbf{u}_h^{(i)2}\right) + \sum_{j=1}^{h} \frac{\partial}{\partial \mathbf{x}^{(j)}} \cdot \left(\int \cdots \int f_h \xi^{(j)} \tfrac{1}{2} \xi^{(i)2} \prod_{k=1}^{h} d\xi^{(k)}\right)$$

$$= n_h \mathbf{u}_h^{(i)} \eta_h^{(i)} - \frac{1}{m} \int n_{h+1} \frac{\partial \phi^{(ih+1)}}{\partial \mathbf{x}^{(i)}} \cdot (\mathbf{u}_{h+1}^{(i)} - \mathbf{u}_h^{(i)}) d\mathbf{x}^{(h+1)}, \tag{2·13}$$

where

$$\frac{3k}{2m} n_h T_h^{(i)} = \int \cdots \int f_h \tfrac{1}{2} \mathbf{v}_h^{(i)2} \prod_{j=1}^{h} d\xi^{(j)}, \tag{2·14}$$

defining a generalized temperature $T_h^{(i)}$, which reduces to the ordinary temperature $T_1^{(i)}$ at the point $\mathbf{x}^{(i)}$ when $h = 1$. Equation (2·13) is first transformed by means of (2·5) and (2·12), with the result

$$\frac{3k}{2m} n_h \frac{dT_h^{(i)}}{dt} + \sum_{j=1}^{h} \frac{\partial}{\partial \mathbf{x}^{(j)}} \cdot \left(\int \cdots \int f_h \mathbf{v}_h^{(j)} \tfrac{1}{2} \xi^{(i)2} \prod_{k=1}^{h} d\xi^{(k)}\right)$$

$$= \frac{1}{m} \mathbf{u}_h^{(i)} \cdot \sum_{j=1}^{h} \frac{\partial}{\partial \mathbf{x}^{(j)}} \cdot \mathsf{k}_h^{(ij)} - \frac{1}{m} \int n_{h+1} \frac{\partial \phi^{(ih+1)}}{\partial \mathbf{x}^{(i)}} \cdot (\mathbf{u}_{h+1}^{(i)} - \mathbf{u}_h^{(i)}) d\mathbf{x}^{(h+1)}. \tag{2·15}$$

Finally, by using (2·8), the following is obtained.

$$\tfrac{3}{2}kn_h\frac{dT_h^{(i)}}{dt}+\sum_{j=1}^{h}\left\{\frac{\partial}{\partial\mathbf{x}^{(j)}}\cdot\mathbf{m}_h^{(ij)}+\left(\mathbf{k}_h^{(ij)}\cdot\frac{\partial}{\partial\mathbf{x}^{(j)}}\right)\mathbf{u}_h^{(i)}\right\}=-\int n_{h+1}\frac{\partial\phi^{(ih+1)}}{\partial\mathbf{x}^{(i)}}\,(\mathbf{u}_{h+1}^{(i)}-\mathbf{u}_h^{(i)})\,d\mathbf{x}^{(h+1)},\tag{2·16}$$

where

$$\mathbf{m}_h^{(ij)}=m\int^{(h)}\!\!\int f_h\mathbf{v}_h^{(i)}\tfrac{1}{2}\mathbf{v}^{(j)2}\prod_{k=1}^{h}d\boldsymbol{\xi}^{(k)}.\tag{2·17}$$

(2·16) is not yet the energy equation, as it refers only to the thermal energy of motion $\frac{3}{2}kT_h^{(i)}$. The corresponding equation for the potential energy is derived by writing

$$I_h^{(i)}=\frac{1}{2}\sum_{j=1}^{h}\phi^{(ij)}+\frac{1}{2}\int\frac{n_{h+1}}{n_h}\phi^{(ih+1)}d\mathbf{x}^{(h+1)}\tag{2·18}$$

for I in (2·5), note that a factor $\frac{1}{2}$ must be inserted in localizing the potential energy, to show that the energy is shared by the two molecules concerned, otherwise an incorrect expression for the internal energy would result from integrating over the whole fluid. It is found that

$$n_h\frac{dI_h^{(i)}}{dt}=\frac{\partial}{\partial t}\left\{\tfrac{1}{2}n_h\sum_{j=1}^{h}\phi^{(ij)}+\frac{1}{2}\int n_{h+1}\phi^{(ih+1)}d\mathbf{x}^{(h+1)}\right\}+\sum_{j=1}^{h}\frac{\partial}{\partial\mathbf{x}^{(j)}}\cdot(n_h I_h^{(i)}\mathbf{u}_h^{(j)}),\tag{2·19}$$

and so, with the aid of (2·3),

$$\begin{aligned}n_h\frac{dI_h^{(i)}}{dt}&+\sum_{j=1}^{h}\frac{\partial}{\partial\mathbf{x}^{(j)}}\left\{\frac{1}{2}\int n_{h+1}\phi^{(ih+1)}(\mathbf{u}_{h+1}^{(j)}-\mathbf{u}_h^{(j)})\,d\mathbf{x}^{(h+1)}\right\}\\&=-\tfrac{1}{2}n_h\sum_{j=1}^{h}(\mathbf{u}_h^{(j)}-\mathbf{u}_h^{(i)})\cdot\frac{\partial\phi^{(ij)}}{\partial\mathbf{x}^{(i)}}-\frac{1}{2}\int n_{h+1}(\mathbf{u}_{h+1}^{(h+1)}-\mathbf{u}_{h+1}^{(i)})\,\frac{\partial\phi^{(ih+1)}}{\partial\mathbf{x}^{(i)}}\,d\mathbf{x}^{(h+1)}.\end{aligned}\tag{2·20}$$

This may be written in the form

$$n_h\frac{dI_h^{(i)}}{dt}+\sum_{j=1}^{h}\frac{\partial}{\partial\mathbf{x}^{(j)}}\,\mathbf{n}_h^{(ij)}=\int n_{h+1}\frac{\partial\phi^{(ih+1)}}{\partial\mathbf{x}^{(i)}}\cdot(\mathbf{u}_{h+1}^{(i)}-\mathbf{u}_h^{(i)})\,d\mathbf{x}^{(h+1)},\tag{2·21}$$

where

$$\begin{aligned}\sum_{j=1}^{h}\frac{\partial}{\partial\mathbf{x}^{(j)}}\cdot\mathbf{n}_h^{(ij)}&=\sum_{j=1}^{h}\frac{\partial}{\partial\mathbf{x}^{(j)}}\left\{\frac{1}{2}\int n_{h+1}\phi^{(ih+1)}(\mathbf{u}_{h+1}^{(j)}-\mathbf{u}_h^{(j)})\,d\mathbf{x}^{(h+1)}\right\}\\&+\tfrac{1}{2}n_h\sum_{j=1}^{h}(\mathbf{u}_h^{(j)}-\mathbf{u}_h^{(i)})\cdot\frac{\partial\phi^{(ij)}}{\partial\mathbf{x}^{(i)}}+\frac{1}{2}\int n_{h+1}(\mathbf{u}_{h+1}^{(h+1)}+\mathbf{u}_{h+1}^{(i)}-2\mathbf{u}_h^{(i)})\,\frac{\partial\phi^{(ih+1)}}{\partial\mathbf{x}^{(i)}}\,d\mathbf{x}^{(h+1)}.\end{aligned}\tag{2·22}$$

Then, on adding (2·16) and (2·22), the following composite equation results:

$$n_h\frac{dE_h^{(i)}}{dt}+\sum_{j=1}^{h}\left\{\frac{\partial}{\partial\mathbf{x}^{(j)}}\,\mathbf{q}_h^{(ij)}+\left(\mathbf{k}_h^{(ij)}\cdot\frac{\partial}{\partial\mathbf{x}^{(j)}}\right)\mathbf{u}_h^{(i)}\right\}=0,\tag{2·23}$$

where

$$E_h^{(i)}=\tfrac{3}{2}kT_h^{(i)}+I_h^{(i)},\tag{2·24}$$

and

$$\mathbf{q}_h^{(ij)}=\mathbf{m}_h^{(ij)}+\mathbf{n}_h^{(ij)}.\tag{2·25}$$

(2·23) may now be regarded as the generalization of the equation of energy transport, and it is seen that the energy flux $\mathbf{q}_h^{(ij)}$, like the pressure tensor, must be regarded as having two parts, due to the thermal motion and the molecular forces respectively.

3 The general expansion procedure

It was pointed out in § 2 of part II that the integro-differential equation (2·1) possesses solutions much more general than those which, apart from minute random deviations, describe the condition of actual fluids, the latter we call 'normal' solutions

A normal solution is determined completely when the density n_1, the mean velocity $\mathbf{u}_1$, the temperature T_1, and the external force $\mathbf{P}$, are known throughout the fluid; it is sufficient to know the first three of these at a given time t, since the hydrodynamical equations (2·3), (2·6) and (2·13) for $h = 1$ determine their subsequent variation. The more general solutions are distinguished from the normal solutions in that they may depend on the time t directly, as well as indirectly through the density, velocity, and temperature, also they may depend on the co-ordinates explicitly in a more general way than we expect of the normal solutions, namely, as a symmetrical function of the relative co-ordinates of the h molecules As only the normal solutions of (2·1) are required for the present purpose, we postulate

$$f_h = F_h\left(\mathbf{x}^{(i)}, \boldsymbol{\xi}^{(i)}, \lambda_k, \frac{\partial \lambda_k}{\partial \mathbf{x}}, \frac{\partial^2 \lambda_k}{\partial \mathbf{x}\,\partial \mathbf{x}}, \ldots\right), \tag{3·1}$$

where the $\mathbf{x}^{(i)}$ occur only symmetrically in the combinations $\mathbf{r}^{(ij)} = \mathbf{x}^{(j)} - \mathbf{x}^{(i)}$, and the λ_k ($k = 1, \ldots, 8$) represent the values n, $\mathbf{u}$, T and $\mathbf{P}$ of the density, velocity, temperature, and force respectively at the *mean centre* $\mathbf{x} = \frac{1}{h}\sum_{i=1}^{h} \mathbf{x}^{(i)}$ of the molecules. From the definitions of n_h, $\mathbf{u}_h^{(i)}$, $T_h^{(i)}$, $\mathbf{v}_h^{(i)}$, $\mathbf{p}_h^{(ij)}$, $\mathbf{q}_h^{(ij)}$, etc., it then follows that all these quantities will be functions of the same kind as F_h; for example,

$$n_h = N_h\left(\mathbf{x}^{(i)}, \boldsymbol{\xi}^{(i)}, \lambda_k, \frac{\partial \lambda_k}{\partial \mathbf{x}}, \ldots\right),$$

where again the $\mathbf{x}^{(i)}$ occur only in the combinations $\mathbf{r}^{(ij)}$, so that $\sum_{i=1}^{h} \frac{\partial N_h}{\partial \mathbf{x}^{(i)}} = 0$. More generally, if c represents any of the quantities enumerated,

$$c = C\left(\mathbf{x}^{(i)}, \boldsymbol{\xi}^{(i)}, \lambda_k, \frac{\partial \lambda_k}{\partial \mathbf{x}}, \frac{\partial^2 \lambda_k}{\partial \mathbf{x}\,\partial \mathbf{x}}, \ldots\right), \tag{3·2}$$

the small letter always being replaced by the corresponding large letter to denote explicit dependence on the λ_k.

It is now supposed that C is expanded in powers of $\frac{\partial \lambda_k}{\partial \mathbf{x}}, \frac{\partial^2 \lambda_k}{\partial \mathbf{x}\,\partial \mathbf{x}}, \ldots$, so that

$$c = c^0 + c' + c'' + \ldots, \tag{3·3}$$

where
$$\left.\begin{aligned} c^0 &= C^0(\mathbf{x}^{(i)}, \boldsymbol{\xi}^{(i)}, \lambda_k), \\ c' &= \sum_k \frac{\partial \lambda_k}{\partial \mathbf{x}} \cdot \mathbf{C}'_k(\mathbf{x}^{(i)}, \boldsymbol{\xi}^{(i)}, \lambda_k), \\ c'' &= \sum_k \frac{\partial^2 \lambda_k}{\partial \mathbf{x} \partial \mathbf{x}} \cdot \mathbf{C}''_k(\mathbf{x}^{(i)}, \boldsymbol{\xi}^{(i)}, \lambda_k) + \sum_{k,l} \left(\frac{\partial \lambda_k}{\partial \mathbf{x}} \frac{\partial \lambda_l}{\partial \mathbf{x}}\right) : \mathbf{C}''_{kl}(\mathbf{x}^{(i)}, \boldsymbol{\xi}^{(i)}, \lambda_k), \end{aligned}\right\} \tag{3·4}$$

and, in general, $c^{(r)}$ involves only r derivatives of the λ_k.

From the hydrodynamical equations (2·3), (2·12) and (2·16) with $h = 1$, it follows that

$$\begin{aligned} \frac{\partial c^0}{\partial t} &= \sum_k \frac{\partial C^0}{\partial \lambda_k} \frac{\partial \lambda_k}{\partial t} \\ &= -\frac{\partial C^0}{\partial n}\left\{\frac{\partial}{\partial \mathbf{x}}(n\mathbf{u})\right\} - \frac{\partial C^0}{\partial \mathbf{u}}\left(\mathbf{u} \cdot \frac{\partial}{\partial \mathbf{x}}\mathbf{u} + \frac{1}{mn}\frac{\partial}{\partial \mathbf{x}} \cdot \mathbf{p} - \frac{1}{m}\mathbf{P}\right) \\ &\quad - \frac{\partial C^0}{\partial T}\left\{\mathbf{u}\,\frac{\partial T}{\partial \mathbf{x}} + \frac{2}{3kn}\left(\frac{\partial}{\partial \mathbf{x}}\,\mathbf{m} + \mathbf{k} \cdot \frac{\partial}{\partial \mathbf{x}} \cdot \mathbf{u} + \int \frac{\partial \phi}{\partial \mathbf{x}^{(1)}} \cdot (\mathbf{u}_2^{(1)} - \mathbf{u})\, n_2\, d\mathbf{x}^{(2)}\right)\right\} + \frac{\partial C^0}{\partial \mathbf{P}} \cdot \frac{\partial}{\partial t}\mathbf{P}. \end{aligned} \tag{3·5}$$

Also
$$\frac{\partial c'}{\partial t} = \sum_k \frac{\partial C'}{\partial \lambda_k}\frac{\partial \lambda_k}{\partial t} + \sum_k \mathbf{C}'_k \cdot \frac{\partial}{\partial \mathbf{x}}\frac{\partial \lambda_k}{\partial t}, \tag{3·6}$$

whilst
$$\frac{\partial c^0}{\partial \mathbf{x}^{(i)}} = \frac{\partial C^0}{\partial \mathbf{x}^{(i)}} + \frac{1}{h}\sum_k \frac{\partial \lambda_k}{\partial \mathbf{x}}\frac{\partial C^0}{\partial \lambda_k}, \tag{3·7}$$

and
$$\frac{\partial c'}{\partial \mathbf{x}^{(i)}} = \frac{\partial C'}{\partial \mathbf{x}^{(i)}} + \frac{1}{h}\sum_k \frac{\partial \lambda_k}{\partial \mathbf{x}}\frac{\partial C'}{\partial \lambda_k} + \frac{1}{h}\sum_k \mathbf{C}'_k \frac{\partial}{\partial \mathbf{x}}\frac{\partial \lambda_k}{\partial \mathbf{x}}. \tag{3·8}$$

It must be noted that in calculating integrals like $\int \frac{\partial \phi^{(ih+1)}}{\partial \mathbf{x}^{(i)}} c_{h+1}\, d\mathbf{x}^{(h+1)}$, where c_{h+1} is a function differing from $c = c_h$ only in containing $h+1$ instead of h co-ordinates and velocities, account must be taken of the shift of the mean centre from

$$\mathbf{x}_{h+1} = \frac{1}{h+1}\sum_{i=1}^{h+1} \mathbf{x}^{(i)} \quad \text{to} \quad \mathbf{x} = \mathbf{x}_h = \frac{1}{h}\sum_{i=1}^{h} \mathbf{x}^{(i)}.$$

Thus c_{h+1} must be replaced by

$$c_{h+1} + \sum_k \frac{\partial c_{h+1}}{\partial \lambda_k}\frac{\partial \lambda_k}{\partial \mathbf{x}} \cdot (\mathbf{x}_{h+1} - \mathbf{x}) + \ldots$$

when $\lambda_k(\mathbf{x}_{h+1})$ is replaced by $\lambda_k(\mathbf{x})$, the actual displacement $\mathbf{d}$ of the mean centre is

$$\mathbf{d} = \mathbf{x}_{h+1} - \mathbf{x} = \frac{1}{h(h+1)}\sum_{i=1}^{h} (\mathbf{x}^{(h+1)} - \mathbf{x}^{(i)}). \tag{3·9}$$

Hence

$$\int \frac{\partial \phi^{(ih+1)}}{\partial \mathbf{x}^{(i)}} c_{h+1}\, d\mathbf{x}^{(h+1)} \to \int \frac{\partial \phi^{(ih+1)}}{\partial \mathbf{x}^{(i)}}\left(c_{h+1} + \mathbf{d} \cdot \frac{\partial c_{h+1}}{\partial \mathbf{x}} + \tfrac{1}{2}(\mathbf{dd}) : \frac{\partial^2 c_{h+1}}{\partial \mathbf{x}\partial \mathbf{x}} + \ldots\right) d\mathbf{x}^{(h+1)}, \tag{3·10}$$

where $\frac{\partial}{\partial \mathbf{x}}$ signifies $\sum_k \frac{\partial \lambda_k}{\partial \mathbf{x}}\frac{\partial}{\partial \lambda_k} + \sum_k \frac{\partial^2 \lambda_k}{\partial \mathbf{x}\partial \mathbf{x}} \frac{\partial}{\partial\left(\frac{\partial \lambda_k}{\partial \mathbf{x}}\right)} + \ldots$.

When $f_h^0, f_h', \ldots$ are known, it is easy to calculate $c^0, c', \ldots$ explicitly, for example,

$$\left.\begin{aligned} n_h^0 &= \int^{(h)} \ldots \int f_h^0 \prod_i d\boldsymbol{\xi}^{(i)}, \\ n_h' &= \int^{(h)} \ldots \int f_h' \prod_i d\boldsymbol{\xi}^{(i)}, \end{aligned}\right\} \tag{3·11}$$

$$\left.\begin{aligned} \mathbf{u}_h^{(i)0} &= \int^{(h)} \ldots \int \frac{f_h^0}{n_h^0} \boldsymbol{\xi}^{(i)} \prod_j d\boldsymbol{\xi}^{(j)}, \quad \mathbf{v}_h^{(i)0} = \boldsymbol{\xi}^{(i)} - \mathbf{u}_h^{(i)0}, \\ \mathbf{u}_h^{(i)'} &= \int^{(h)} \ldots \int \frac{f_h'}{n_h^0} \mathbf{v}_h^{(i)0} \prod_j d\boldsymbol{\xi}^{(j)}, \quad \mathbf{v}_h^{(i)'} = -\mathbf{u}_h^{(i)'}, \end{aligned}\right\} \tag{3·12}$$

$$\left.\begin{aligned} T_h^{(i)0} &= \frac{2m}{3k} \int^{(h)} \int \frac{f_h^0}{n_h^0} \tfrac{1}{2} \mathbf{v}_h^{(i)02} \prod_j d\boldsymbol{\xi}^{(j)}, \\ T_h^{(i)'} &= \frac{2m}{3k} \int^{(h)} \int \frac{f_h'}{n_h^0} \left(\tfrac{1}{2} \mathbf{v}_h^{(i)02} - \frac{3k}{2m} T_h^{(i)0} \right) \prod_j d\boldsymbol{\xi}^{(j)}, \end{aligned}\right\} \tag{3·13}$$

$$\left.\begin{aligned} \mathbf{k}_h^{(ij)0} &= m \int^{(h)} \ldots \int f_h^0 \mathbf{v}_h^{(i)0} \mathbf{v}_h^{(j)0} \prod_k d\boldsymbol{\xi}^{(k)}, \\ \mathbf{k}_h^{(ij)'} &= m \int^{(h)} \ldots \int f_h' \mathbf{v}_h^{(i)0} \mathbf{v}_h^{(j)0} \prod_k d\boldsymbol{\xi}^{(k)}, \end{aligned}\right\} \tag{3·14}$$

$$\left.\begin{aligned} \mathbf{m}_h^{(ij)0} &= m \int^{(h)} \ldots \int f_h^0 \mathbf{v}_h^{(i)0} \tfrac{1}{2} \mathbf{v}_h^{(j)02} \prod_k d\boldsymbol{\xi}^{(k)}, \\ \mathbf{m}_h^{(ij)'} &= m \int^{(h)} \ldots \int f_h' \mathbf{v}_h^{(i)0} \left(\tfrac{1}{2} \mathbf{v}_h^{(j)02} - (\tfrac{3}{2} + \delta_{ij}) \frac{k}{m} T_h^{(j)0} \right) \prod_k d\boldsymbol{\xi}^{(k)}, \end{aligned}\right\} \tag{3·15}$$

$$\left.\begin{aligned} m\boldsymbol{\eta}_h^{(i)0} &= -\sum_j \frac{\partial \phi^{(ij)}}{\partial \mathbf{x}^{(i)}} - \int \frac{\partial \phi^{(ih+1)}}{\partial \mathbf{x}^{(i)}} \frac{n_{h+1}^0}{n_h^0} d\mathbf{x}^{(h+1)} + \mathbf{P}, \\ m\boldsymbol{\eta}_h^{(i)'} &= -\int \frac{\partial \phi^{(ih+1)}}{\partial \mathbf{x}^{(i)}} \left(\frac{n_{h+1}}{n_h} \right)' d\mathbf{x}^{(h+1)} - \frac{1}{n_h^0} \int \mathbf{d} \cdot \sum_k \frac{\partial \lambda_k}{\partial \mathbf{x}} \frac{\partial N_{h+1}^0}{\partial \lambda_k} \frac{\partial \phi^{(ih+1)}}{\partial \mathbf{x}^{(i)}} d\mathbf{x}^{(h+1)} \\ &\qquad + (\mathbf{x}^{(i)} - \mathbf{x}) \cdot \frac{\partial}{\partial \mathbf{x}} \mathbf{P}. \end{aligned}\right\} \tag{3·16}$$

One further remark may be made concerning the dependence of F_h on $\mathbf{u}$ it is clear, on physical grounds, that the distribution of velocities among the molecules of a fluid about the mean $\mathbf{u}$ is independent of any translational motion of the fluid as a whole. Hence F_h will contain $\mathbf{u}$ only in the combinations $\mathbf{v}^{(i)} = \boldsymbol{\xi}^{(i)} - \mathbf{u}$, and in the derivatives $\frac{\partial}{\partial \mathbf{x}} \mathbf{u}$, $\frac{\partial^2}{\partial \mathbf{x} \partial \mathbf{x}} \mathbf{u}$, etc., so that

$$\frac{\partial F_h}{\partial \mathbf{u}} + \sum_{i=1}^{h} \frac{\partial F_h}{\partial \boldsymbol{\xi}^{(i)}} = 0. \tag{3·17}$$

By integration over the velocities it follows also that

$$\frac{\partial N_h}{\partial \mathbf{u}} = 0. \tag{3·18}$$

The equation to determine F_h^0 is easily derived from (2·1), it follows from (3·5) that $\left(\frac{\partial f_h}{\partial t}\right)^0 = \frac{1}{m}\mathbf{P}\,\frac{\partial F_h^0}{\partial \mathbf{u}}$, so

$$\sum_{i=1}^{h}\left(m\boldsymbol{\xi}^{(i)}\,\frac{\partial F_h^0}{\partial \mathbf{x}^{(i)}} - \sum_{j=1}^{h}\frac{\partial \phi^{(ij)}}{\partial \mathbf{x}^{(i)}}\cdot\frac{\partial F_h^0}{\partial \boldsymbol{\xi}^{(i)}}\right) = \sum_{i=1}^{h}\int\int\frac{\partial \phi^{(ih+1)}}{\partial \mathbf{x}^{(i)}}\cdot\frac{\partial F_{h+1}^0}{\partial \boldsymbol{\xi}^{(i)}}\,d\mathbf{x}^{(h+1)}d\boldsymbol{\xi}^{(h+1)} \tag{3·19}$$

It is unnecessary to solve this equation, however, as it has been obtained from (2·1) by writing $\partial\lambda_k/\partial\mathbf{x} = 0$, i.e. by postulating equilibrium conditions, and the solution of (3·19) must therefore coincide with the known equilibrium solution of (2·1):

$$F_h^0 = N_h^0\left(\frac{m}{2\pi kT}\right)^{\frac{3}{2}h}\exp\left(-\frac{m}{2kT}\sum_{i=1}^{h}\mathbf{v}^{(i)2}\right). \tag{3·20}$$

Here, in accordance with (3·18), N_h^0 does not contain $\mathbf{u}$, though it is, in general, a function of n and T as well as the relative co-ordinates. Substitution of (3·20) in (3·11) to (3·15) now leads to $n_1^0 = n$, and

$$\left.\begin{aligned}&\mathbf{u}_h^{(i)0} = \mathbf{u}, \quad \mathbf{v}_h^{(i)0} = \mathbf{v}^{(i)} = \boldsymbol{\xi}^{(i)} - \mathbf{u},\\ &T_h^{(i)0} = T, \quad \mathbf{k}_h^{(ij)0} = N_h^0 kT\delta_{ij}\mathbf{1}, \quad \mathbf{m}_h^{(ij)0} = 0,\end{aligned}\right\} \tag{3·21}$$

it follows from these that n_1', T_1', $\mathbf{u}_1'$, n_1'', T_1'', $\mathbf{u}_1''$, etc., must all vanish.

From (3·5) it is found that

$$\left(\frac{\partial n_h}{\partial t}\right)^0 = 0, \quad \left(\frac{\partial}{\partial t}\mathbf{u}_h^{(i)}\right)^0 = \frac{1}{m}\mathbf{P}, \quad \left(\frac{\partial T_h^{(i)}}{\partial t}\right)^0 = 0, \tag{3·22}$$

and

$$\left(\frac{\partial n_h}{\partial t}\right)' = -\frac{\partial N_h^0}{\partial n}\left(\frac{\partial}{\partial \mathbf{x}}\cdot(n\mathbf{u})\right) - \frac{\partial N_h^0}{\partial T}\left(\mathbf{u}\,\frac{\partial T}{\partial \mathbf{x}} + \tfrac{2}{3}T\frac{\partial}{\partial \mathbf{x}}\cdot\mathbf{u} + \frac{2}{3kn}\int\frac{\partial \phi}{\partial \mathbf{x}^{(1)}}\,\mathbf{u}_2^{(1)\prime}n_2^0\,d\mathbf{x}^{(2)}\right). \tag{3·23}$$

These expressions can also be calculated from the hydrodynamical equations (2·3), (2·6) and (2·16), which give

$$\left(\frac{\partial n_h}{\partial t}\right)^0 = 0, \quad \left(\frac{\partial}{\partial t}\mathbf{u}_h^{(i)}\right)^0 = \frac{1}{m}\left(-\frac{kT}{N_h^0}\frac{\partial N_h^0}{\partial \mathbf{x}^{(i)}} + \boldsymbol{\eta}_h^{(i)0}\right), \quad \left(\frac{\partial T_h^{(i)}}{\partial t}\right)^0 = 0, \tag{3·24}$$

and

$$\left(\frac{\partial n_h}{\partial t}\right)' = -\left\{N_h^0\frac{\partial}{\partial \mathbf{x}}\,\mathbf{u} + \frac{\partial N_h^0}{\partial n}\mathbf{u}\cdot\frac{\partial n}{\partial \mathbf{x}} + \frac{\partial N_h^0}{\partial T}\mathbf{u}\cdot\frac{\partial T}{\partial \mathbf{x}} + \sum_{i=1}^{h}\frac{\partial}{\partial \mathbf{x}^{(i)}}\cdot(N_h^0\mathbf{u}_h^{(i)\prime})\right\}, \tag{3·25}$$

since $\sum_{i=1}^{h}\frac{\partial N_h'}{\partial \mathbf{x}^{(i)}}$ vanishes.

Comparison of the second of equations (3·22) and (3·24), with reference to (3·16), leads to

$$\frac{kT}{N_h^0}\frac{\partial N_h^0}{\partial \mathbf{x}^{(i)}} + \sum_{j=1}^{h}\frac{\partial \phi^{(ij)}}{\partial \mathbf{x}^{(i)}} + \int\frac{N_{h+1}^0}{N_h^0}\frac{\partial \phi^{(ih+1)}}{\partial \mathbf{x}^{(i)}}\,d\mathbf{x}^{(h+1)} = 0, \tag{3·26}$$

which agrees precisely with the equation, first found in part I, for determining the equilibrium value of n_h, thus confirming our adoption of the equilibrium form of f_h for F_h^0. Note that the external force $\mathbf{P}(\mathbf{x}^{(i)})$ does not influence the 'zero' approximation in any explicit way, this could be foreseen from the fact that the constant force $\mathbf{P}(\mathbf{x})$ accelerates all the molecules in the same way, and can therefore influence only the density distribution, which has been assumed to be given.

The following equation is found by comparing (3·23) and (3·25); and writing $h = 2$ and $\mathbf{r} = \mathbf{x}^{(2)} - \mathbf{x}^{(1)}$:

$$\frac{\partial}{\partial \mathbf{r}}.\{N_2^0(\mathbf{U}_2^{(2)\prime} - \mathbf{U}_1^{(1)\prime})\} = \left(n\frac{\partial N_2^0}{\partial n} - N_2^0 + \tfrac{3}{2}T\frac{\partial N_2^0}{\partial T}\right)\frac{\partial}{\partial \mathbf{x}}.\mathbf{u} + \frac{1}{3kn}\frac{\partial N_2^0}{\partial T}\int\frac{\partial \phi}{\partial \mathbf{r}}.(\mathbf{U}_2^{(2)\prime} - \mathbf{U}_2^{(1)\prime})\,N_2^0 d\mathbf{r} \tag{3·27}$$

This may be regarded as an integro-differential equation for the (partial) determination of $\mathbf{U}_2^{(2)\prime} - \mathbf{U}_2^{(1)\prime}$. Multiplication by $\phi(r)$ and integration over $\mathbf{r}$ gives

$$\int\frac{\partial \phi}{\partial \mathbf{r}}\,(\mathbf{U}_2^{(2)\prime} - \mathbf{U}_2^{(1)\prime})\,N_2^0 d\mathbf{r} = nkT\sigma\frac{\partial}{\partial \mathbf{x}}.\mathbf{u}, \tag{3·28}$$

where

$$nkT\sigma = -\frac{\int\left(n\frac{\partial N_2^0}{\partial n} - N_2^0 + \frac{3}{2}T\frac{\partial N_2^0}{\partial T}\right)\phi\, d\mathbf{r}}{1 + \frac{1}{3kn}\int\frac{\partial N_2^0}{\partial T}\phi\, d\mathbf{r}}, \tag{3·29}$$

so that

$$N_2^0(\mathbf{U}_2^{(2)\prime} - \mathbf{U}_2^{(1)\prime}) = \frac{\mathbf{r}}{r^3}\int_0^r\left(n\frac{\partial N_2^0}{\partial n} - N_2^0 + \tfrac{1}{2}(\sigma+2)\,T\frac{\partial N_2^0}{\partial T}\right)r^3 dr\,\frac{\partial}{\partial \mathbf{x}}.\mathbf{u} + \mathbf{v}, \tag{3·30}$$

provided $\dfrac{\partial}{\partial \mathbf{r}}\,\mathbf{v} = 0$. It will appear later how $\mathbf{v}$ can be determined from other sources. At present we wish to draw attention to the failure of the above solution when

$$\frac{\partial}{\partial T}\int\frac{N_2^0}{n}\phi\, d\mathbf{r} = -3k, \tag{3·31}$$

i.e. when the internal energy of the fluid is changing at a certain rate with temperature. On approaching this point, the average relative velocity of two molecules at a given separation tends to infinity, unless $\dfrac{\partial}{\partial \mathbf{x}}\,\mathbf{u} = 0$, this means that the condition of the fluid, even in equilibrium, is highly unstable for small changes in the volume. One can hardly fail to associate such a condition with condensation, or rather the limit of superheating above the boiling point, and we therefore regard (3·31) as the exact analogue of the 'condition for condensation' found in part II.

4 Viscosity and thermal conduction in liquids

Before proceeding to calculate the viscosity and thermal conduction coefficients, the form of F_1', N_2' and $N_2^0(\mathbf{U}_2^{(1)\prime} + \mathbf{U}_2^{(2)\prime})$ which can be determined qualitatively from quite general considerations, is required. Write

$$\left.\begin{aligned} \mathbf{a} &= \frac{\partial T}{\partial \mathbf{x}}, \quad \mathbf{a}' = \frac{\partial n}{\partial \mathbf{x}}, \\ b &= \frac{\partial}{\partial \mathbf{x}}.\mathbf{u}, \quad b' = \frac{\partial}{\partial \mathbf{x}}.\mathbf{P}, \\ \mathbf{b} &= \frac{1}{2}\left(\frac{\partial}{\partial \mathbf{x}}\mathbf{u} + \overline{\frac{\partial}{\partial \mathbf{x}}\mathbf{u}}\right) - \frac{1}{3}\frac{\partial}{\partial \mathbf{x}}.\mathbf{u}\mathbf{1}, \\ \mathbf{b}' &= \frac{1}{2}\left(\frac{\partial}{\partial \mathbf{x}}\mathbf{P} + \overline{\frac{\partial}{\partial \mathbf{x}}\mathbf{P}}\right) - \frac{1}{3}\frac{\partial}{\partial \mathbf{x}}.\mathbf{P}\mathbf{1}. \end{aligned}\right\} \tag{4·1}$$

Now F_1' is linear in the a's and b's, can contain besides only n, T and $\mathbf{v}$, and satisfies $\int F_1' d\mathbf{v} = N_1' = 0$, hence it is of the form

$$F_1' = \varphi_1 \mathbf{v}\ \mathbf{a} + \varphi_1' \mathbf{v}\ \mathbf{a}' + \varphi_2 \mathbf{v}.\mathbf{b}.\mathbf{v} + \varphi_2' \mathbf{v}\ \mathbf{b}'\ \mathbf{v}, \tag{4·2}$$

where φ_1, φ_1', φ_2 and φ_2' contain only n, T and v.

N_2' is linear in the a's and b's, can contain besides only n, T and $\mathbf{r}$, it is unaltered on replacing $\mathbf{r}$ by $-\mathbf{r}$ (interchange of $\mathbf{x}^{(1)}$ and $\mathbf{x}^{(2)}$), and satisfies $\int N_2' d\mathbf{r} = (N-1) N_1' = 0$, hence it is of the form

$$N_2' = \nu \mathbf{r}.\mathbf{b}\ \mathbf{r} + \nu' \mathbf{r}.\mathbf{b}'\ \mathbf{r}, \tag{4·3}$$

where ν and ν' are functions of n, T and r only

$N_2^0(\mathbf{U}_2^{(1)'} + \mathbf{U}_2^{(2)'})$ is linear in the a's and b's, can contain besides only n, T and $\mathbf{r}$, is unaltered on replacing $\mathbf{r}$ by $-\mathbf{r}$, and satisfies

$$\int N_2^0 (\mathbf{U}_2^{(1)'} + \mathbf{U}_2^{(2)'}) d\mathbf{r} = 2n(N-1)\mathbf{u}_1' = 0;$$

hence it is of the form

$$N_2^0(\mathbf{U}_2^{(1)'} + \mathbf{U}_2^{(2)'}) = \tau(\mathbf{r}\ \mathbf{a}\mathbf{r} - \tfrac{1}{3} r^2 \mathbf{a}) + \tau'(\mathbf{r}\ \mathbf{a}'\mathbf{r} - \tfrac{1}{3} r^2 \mathbf{a}'), \tag{4·4}$$

where τ and τ' are functions of n, T and r only The form of other functions involving the a's and b's may be determined in a similar way, but here it is necessary to note only that N_2'' and N_2''', like N_2', must be even functions of $\mathbf{r}$.

Now turn to the equation (2·11), which, with $h = 1$, determines the part I of the pressure tensor due to the molecular forces, namely,

$$\frac{\partial}{\partial \mathbf{x}}\ \mathsf{I} = -\int n_2 \frac{\partial \phi}{\partial \mathbf{r}} d\mathbf{x}^{(2)}. \tag{4·5}$$

When each side of this equation is expanded in the way shown in (3 3) and (3·10) (substituting n_2 for c_{h+1} in the latter), the following are obtained:

$$\frac{\partial}{\partial \mathbf{x}}\ \mathsf{I}^0 = -\frac{\partial}{\partial \mathbf{x}} \int \tfrac{1}{2} \mathbf{r} N_2^0 \frac{\partial \phi}{\partial \mathbf{r}} d\mathbf{r}, \tag{4·6}$$

$$\frac{\partial}{\partial \mathbf{x}}\ \mathsf{I}' = -\frac{\partial}{\partial \mathbf{x}} . \int \tfrac{1}{2} \mathbf{r} N_2' \frac{\partial \phi}{\partial \mathbf{r}} d\mathbf{r}, \tag{4·7}$$

$$\frac{\partial}{\partial \mathbf{x}} . \mathsf{I}'' = -\frac{\partial}{\partial \mathbf{x}} . \int \tfrac{1}{2} \mathbf{r} N_2'' \frac{\partial \phi}{\partial \mathbf{r}} d\mathbf{r} - \frac{1}{6} \frac{\partial^3}{\partial \mathbf{x} \partial \mathbf{x} \partial \mathbf{x}} . \int \tfrac{1}{2} \mathbf{r}\, \tfrac{1}{2} \mathbf{r}\, \tfrac{1}{2} \mathbf{r} N_2^0 \frac{\partial \phi}{\partial \mathbf{r}} d\mathbf{r}. \tag{4·8}$$

The other terms on the right-hand side vanish when account is taken of the fact, noted above, that N_2', N_2'' and N_2''' are even functions of $\mathbf{r}$. Then, on integrating (4 6),

$$\begin{aligned} \mathsf{I}^0 &= -\frac{1}{2} \int N_2^0 \mathbf{r} \frac{\partial \phi}{\partial \mathbf{r}} d\mathbf{r} \\ &= -\frac{1}{6} \int N_2^0 \phi'(r) r\, d\mathbf{r}\ \mathsf{1}, \end{aligned} \tag{4·9}$$

in agreement with the formula found previously for the part of the pressure tensor due to the molecular forces in equilibrium. The constant of integration has been rejected to ensure $\mathsf{I}^0 \to 0$ as $n \to 0$.

In the same way, from (4·7),

$$\mathsf{I}' = -\frac{1}{2}\int N_2' \mathbf{r} \frac{\partial \phi}{\partial \mathbf{r}} d\mathbf{r}$$

$$= -\frac{1}{15}\int (\nu \mathbf{b} + \nu' \mathbf{b}') \phi'(r) r^3 d\mathbf{r}, \tag{4·10}$$

on substitution from (4·3) Similar formulae for I'', etc , may be obtained from (4 8), etc To obtain the corresponding expression for $\mathbf{k}'$, substitute (4·2) in (3·14), it is found that

$$\mathbf{k}' = \tfrac{2}{15} m \int (\varphi_2 \mathbf{b} + \varphi_2' \mathbf{b}') v^4 d\mathbf{v}. \tag{4·11}$$

Now the viscosity coefficient μ is defined by the equation

$$\mathbf{p}' = -2\mu \mathbf{b}, \tag{4·12}$$

on the understanding that $\mathbf{b}' = 0$. Hence

$$\mu = \frac{1}{30}\int \nu(r) \phi'(r) r^3 d\mathbf{r} - \frac{1}{15} m \int \varphi_2(v) v^4 d\mathbf{v}. \tag{4 13}$$

In the kinetic theory of gases, no account is taken of the first term in (4·13), due to the molecular forces, the second term is found to be small and to decrease with the temperature. In liquids, however, the viscosity is much larger, and increases with decreasing temperature, it is therefore reasonable to conclude that the first term of (4·13), far from being negligible for the liquid state, is the dominant one in the complete expression. This conclusion is supported by the observation that it is roughly proportional to the square of the density, and so becomes very many times larger in the process of condensation.

To calculate the coefficient of thermal conduction, it is necessary to obtain first the energy flux vector $\mathbf{q}'$. From (2·22) it is found that

$$\frac{\partial}{\partial \mathbf{x}} \mathbf{n}' = -\frac{1}{2}\frac{\partial}{\partial \mathbf{x}} \int \tfrac{1}{2} \mathbf{r} \frac{\partial \phi}{\partial \mathbf{r}} (\mathbf{U}_2^{(1)'} + \mathbf{U}_2^{(2)'}) N_2^0 d\mathbf{r}, \tag{4·14}$$

and, on integration,

$$\mathbf{n}' = -\frac{1}{4}\int \mathbf{r} \frac{\partial \phi}{\partial \mathbf{r}} (\mathbf{U}_2^{(1)'} + \mathbf{U}_2^{(2)'}) N_2^0 d\mathbf{r}$$

$$= -\frac{1}{18}\int (\tau \mathbf{a} + \tau' \mathbf{a}') \phi'(r) r^3 d\mathbf{r}, \tag{4·15}$$

when substitution is made from (4·4). Also, from (3·15),

$$\mathbf{m}' = \frac{1}{6}\int (\varphi_1 \mathbf{a} + \varphi_1' \mathbf{a}') \left(v^2 - \frac{5kT}{m}\right) v^2 d\mathbf{v}. \tag{4·16}$$

Now, the thermal conductivity coefficient λ is defined by

$$\mathbf{q}' = -\lambda \mathbf{a}, \tag{4·17}$$

on the understanding that $\partial p/\partial \mathbf{x} = 0$, $p = k^0 + l^0$, so that, according to (3·21) and (4·9),

$$\mathbf{a}' = \kappa \mathbf{a},$$

$$\kappa = -\frac{n}{T}\frac{1-\dfrac{1}{6nkT}\displaystyle\int T\frac{\partial N_2^0}{\partial T}\phi'(r)\,r\,d\mathbf{r}}{1-\dfrac{1}{6nkT}\displaystyle\int n\frac{\partial N_2^0}{\partial n}\phi'(r)\,r\,d\mathbf{r}}. \tag{4·18}$$

Hence
$$\lambda = \frac{1}{18}\int(\tau+\kappa\tau')\phi'(r)\,r^3\,d\mathbf{r} - \frac{1}{6}\int(\varphi_1+\kappa\varphi_1')\left(v^2-\frac{5kT}{m}\right)v^2\,d\mathbf{v}, \tag{4·19}$$

where, again, the second part is dominant for the gaseous state, but the first term which does not appear in the theory of gases, becomes of major importance for the liquid state.

5 Preliminary discussion of the coefficients of viscosity and thermal conduction

The complete content of the formulae (4·13) and (4·19) can be discussed only after a rigorous determination of the deformation of the radial distribution function by the motion of the fluid However, this requires the solution of a complicated integro-differential equation, to be developed in the next section, which is a considerable task It therefore seems desirable to derive simple approximations to the exact laws, for comparison with experiment.

It will be assumed that the second term, due to the thermal motion, in each of the formulae (4·13) and (4·19), can be completely neglected for the liquid state.

$\nu\mathbf{r}.\mathbf{b}.\mathbf{r}$, which is the second approximation to the radial distribution function in non-equilibrium conditions, must clearly be proportional to the first approximation $N_2^0(r)$, which gives the density dependence. Hence*

$$\nu = \alpha_1 N_2^0(r), \tag{5·1}$$

where α_1 is a factor of dimensions $[l^{-3}t]$ involving only r and kT/m; accordingly

$$\alpha_1 \sim \frac{1}{r}\left(\frac{m}{kT}\right)^{\frac{1}{2}}. \tag{5·2}$$

Substituting (5·1) in (4·13),

$$\mu \sim \frac{1}{30}\left(\frac{m}{kT}\right)^{\frac{1}{2}}\int N_2^0\phi'(r)\,r^3\,dr. \tag{5·3}$$

The factor N_2^0 in the integrand of (5·3) rises from zero to a sharp maximum at the distance r_1 of the nearest neighbours from a given molecule. When account is taken

* A closer study shows that the details of the following calculation require amendment; but the result (5 11) is still unaffected.

of the second factor $\phi'(r)r^4$, this maximum is only slightly displaced, and the integrand is almost zero elsewhere. Hence the integral may be evaluated by the method of 'steepest descent', with the result

$$\mu \sim \frac{2\pi}{15}\left(\frac{m}{kT}\right)^{\frac{1}{2}} N_2^0(r_1)\phi'(r_1)r_1^4\left(\frac{\pi}{\beta}\right)^{\frac{1}{2}}, \tag{5·4}$$

where

$$\beta = \left[\frac{d^2}{dr^2}\{\log N_2^0(r) + \log\phi'(r) + 4\log r\}\right]_{r=r_1}. \tag{5·5}$$

For the further simplification of (5·4), $N_2^0(r)$ may be replaced by its first approximation $n^2 e^{-\phi(r)/kT}$, and, in (5·5), since r_1 is known to lie beyond the potential minimum in the attractive part of the field, it may be assumed that

$$\phi(r) = -\gamma r^{-6}. \tag{5·6}$$

Then

$$\beta = \frac{42\gamma}{kTr_1^8} + \frac{3}{r_1^2} = -\frac{42}{r_1^2}\{\phi(r_1) - \tfrac{1}{14}\}, \tag{5·7}$$

where the second term is generally negligible in comparison with the first Thus

$$\mu \sim \frac{2\pi}{105}(42\pi m\gamma)^{\frac{1}{2}} n^2 r_1 e^{-\phi(r_1)/kT}. \tag{5·8}$$

In order to compare (5·8) with corresponding formulae involving the frequency ν_0 of vibration of the molecules about the equilibrium point r_0 in the molecular field, one may take

$$\phi(r) = -\gamma r^{-6}\{1 - \tfrac{1}{2}(r_0 r^{-1})^6\}, \tag{5·9}$$

according to Lennard-Jones (1924); then ν_0 is given by

$$4\pi^2 m\nu_0^2 = \phi''(r_0) = 36\gamma r_0^{-8}. \tag{5·10}$$

Substituting for γ from (5·10) into (5·8), and taking $n = 2^{-\frac{1}{2}} r_1^{-3}$ (as for a face-centred cubic structure), it is found finally that

$$\mu \sim \frac{\pi^2}{315}(42\pi)^{\frac{1}{2}}\left(\frac{r_0}{r_1}\right)^5 \frac{m\nu_0}{r_0} e^{-\phi(r_1)/kT}. \tag{5·11}$$

Formulae of this type have been suggested before. Andrade (1934) regards viscosity as produced by the transfer of the momentum of molecular vibrations from one 'layer' to a neighbouring one, and obtains, for the melting-point, $\mu \sim m\nu_1/r_1$, where ν_1 is Lindemann's frequency. One recognizes this factor in our formula, but with constant ν_0 referring to the diatomic molecule; the actual dependence on volume is given by the factor $(r_0/r_1)^5 = (v_0/v_1)^{\frac{5}{3}}$. In fact, Andrade's derivation seems to us an ingenious dimensional consideration, which, however, fails to give the correct dependence on volume since dimensionless factors (powers of v_0/v_1) cannot be found in this way. Andrade also finds the exponential factor, but gives no theoretical explanation except that it ought to be of the form of a Boltzmann factor.

A different formula for the coefficient of viscosity has been given by Fürth (1941). It is based on his theory of 'holes'; the liquid is regarded as a continuum containing small bubbles, the size and apparent mass of which depend on the surface tension.

As these bubbles transfer energy and momentum from one layer to another, the coefficient of viscosity can be expressed in terms of the surface tension. Furth's formula contains an exponential factor of the same type as ours; the remaining factor cannot be compared with ours because of the entirely different assumptions made. But Furth has shown that his formula is practically identical with Andrade's because of some empirical relations between the constants involved; and similar considerations can be applied by coupling Fürth's formula with our own, or that of Ewell & Eyring (1937) with our own.

The thermal conductivity may be discussed in a similar way, starting from the rigorous equation (4·19), and omitting the second term as unimportant for the liquid. If one writes

$$\tau = \epsilon N_2^0, \tag{5·12}$$

it follows from simple dimensional considerations that

$$\epsilon \sim \frac{k}{m}\frac{1}{\tau}\left(\frac{m}{kT}\right)^{\frac{1}{2}}. \tag{5·13}$$

It is now seen by substitution in (4·19) that the thermal conductivity consists of two terms, the first of which is proportional to the viscosity, and the second of which depends on the coefficient of expansion, through the parameter κ. Owing to the complication of this second term, it is more difficult to derive a simple formula of general applicability for the thermal conductivity, the complete discussion of which is therefore deferred until a rigorous treatment can be given.

6. The integral equation for F'_h

The equation for the determination of f'_h will now be derived from (2·1) in accordance with the general expansion procedure outlined in § 3. The expression for $\left(\frac{\partial f_h}{\partial t}\right)' = \left(\frac{\partial f_h^0}{\partial t}\right)' + \left(\frac{\partial f'_h}{\partial t}\right)'$ is found to be

$$\begin{aligned}\left(\frac{\partial f_h}{\partial t}\right)' = &-\frac{f_h^0}{n_h^0}\frac{\partial N_h^0}{\partial n}\left\{\frac{\partial}{\partial \mathbf{x}}.(n\mathbf{u})\right\} - \frac{mf_h^0}{kT}\left(\sum_i \mathbf{v}^{(i)}\right).\left\{\mathbf{u}.\frac{\partial}{\partial \mathbf{x}}\mathbf{u} + \frac{1}{mn}\frac{\partial p^0}{\partial \mathbf{x}}\right\}\\ &-\frac{f_h^0}{T}\left\{\frac{T}{n_h^0}\frac{\partial N_h^0}{\partial T} + \frac{m}{2kT}\left(\sum_i v^{(i)2}\right) - \frac{3h}{2}\right\}\left\{\mathbf{u}\,\frac{\partial T}{\partial \mathbf{x}} + \frac{2T}{3}\frac{\partial}{\partial \mathbf{x}}.\mathbf{u} + \frac{\sigma T}{3}\frac{\partial}{\partial \mathbf{x}}\,\mathbf{u}\right\}\\ &+\frac{1}{m}\mathbf{P}.\frac{\partial F'_h}{\partial \mathbf{u}},\end{aligned} \tag{6·1}$$

by substituting f_h for c, and picking out the 'first-order' terms in (3·5) and (3·6). Also, from (3·7) and (3·8),

$$\begin{aligned}\left(\sum_i \boldsymbol{\xi}^{(i)}.\frac{\partial f_h}{\partial \mathbf{x}^{(i)}}\right)' = &\left\{\frac{1}{\bar{h}}\left(\sum_i \mathbf{v}^{(i)}\right) + \mathbf{u}\right\}.\left[\frac{f_h^0}{n_h^0}\frac{\partial N_h^0}{\partial n}\frac{\partial n}{\partial \mathbf{x}} + \frac{mf_h^0}{kT}\left(\sum_i \mathbf{v}^{(i)}\right)\overline{\frac{\partial}{\partial \mathbf{x}}\mathbf{u}}\right.\\ &\left.+\frac{f_h^0}{T}\left\{\frac{T}{n_h^0}\frac{\partial N_h^0}{\partial T} + \frac{m}{2kT}\left(\sum_i v^{(i)2}\right) - \frac{3h}{2}\right\}\frac{\partial T}{\partial \mathbf{x}}\right] + \sum_i \mathbf{v}^{(i)}\frac{\partial F'_h}{\partial \mathbf{x}^{(i)}},\end{aligned} \tag{6·2}$$

and $$\left\{\frac{1}{m}\sum_i\left(\mathbf{P}^{(i)}-\sum_j\frac{\partial\phi^{(ij)}}{\partial\mathbf{x}^{(i)}}\right)\cdot\frac{\partial f_h}{\partial\boldsymbol{\xi}^{(i)}}\right\}'=-\frac{f_h^0}{kT}\sum_i\mathbf{v}^{(i)}\cdot\left\{(\mathbf{x}^{(i)}-\mathbf{x})\cdot\frac{\partial}{\partial\mathbf{x}}\mathbf{P}\right\}$$
$$+\frac{1}{m}\sum_i\left(\mathbf{P}-\sum_j\frac{\partial\phi^{(ij)}}{\partial\mathbf{x}^{(i)}}\right)\cdot\frac{\partial F_h'}{\partial\boldsymbol{\xi}^{(i)}}. \qquad (6\cdot3)$$

(2·1) therefore yields the following equation:

$$-\left[\frac{n}{n_h^0}\frac{\partial N_h^0}{\partial n}+\frac{\sigma+2}{3}\left\{\frac{\partial N_h^0}{\partial T}\frac{T}{n_h^0}+\frac{m}{2kT}\left(\sum_i v^{(i)2}\right)-\frac{3h}{2}\right\}-\frac{m}{3hkT}\left(\sum_i\mathbf{v}^{(i)}\right)^2\right]f_h^0 b$$
$$-\frac{1}{3kT}\sum_i\mathbf{v}^{(i)}.(\mathbf{x}^{(i)}-\mathbf{x})f_h^0 b'+\left\{\frac{\partial N_h^0}{\partial T}\frac{T}{hn_h^0}-\frac{1}{nk}\frac{\partial p^0}{\partial T}+\frac{m}{2hkT}\left(\sum_i v^{(i)2}\right)-\frac{3h}{2}\right\}\frac{f_h^0}{T}\left(\sum_i\mathbf{v}^{(i)}\right).\mathbf{a}$$
$$+\left(\frac{n}{hn_h^0}\frac{\partial N_h^0}{\partial n}-\frac{1}{kT}\frac{\partial p^0}{\partial n}\right)\frac{f_h^0}{n}\left(\sum_i\mathbf{v}^{(i)}\right)\mathbf{a}'+\frac{mf_h^0}{hkT}\left(\sum_i\mathbf{v}^{(i)}\right).\mathbf{b}.\left(\sum_i\mathbf{v}^{(i)}\right)-\frac{f_h^0}{kT}\sum_i\mathbf{v}^{(i)}.\mathbf{b}'.(\mathbf{x}^{(i)}-\mathbf{x})$$
$$-\frac{f_h^0}{2kT}\left\{\sum_i(\mathbf{x}^{(i)}-\mathbf{x})\times\mathbf{v}^{(i)}\right\}.\frac{\partial}{\partial\mathbf{x}}\times\mathbf{P}+\sum_i\left(\mathbf{v}^{(i)}.\frac{\partial F_h'}{\partial\mathbf{x}^{(i)}}-\frac{1}{m}\sum_j\frac{\partial\phi^{(ij)}}{\partial\mathbf{x}^{(i)}}\frac{\partial F_h'}{\partial\boldsymbol{\xi}^{(i)}}\right)$$
$$=\frac{1}{m}\iint\sum_i\frac{\partial\phi^{(ih+1)}}{\partial\mathbf{x}^{(i)}}.\frac{\partial F_{h+1}'}{\partial\boldsymbol{\xi}^{(i)}}d\mathbf{x}^{(h+1)}d\boldsymbol{\xi}^{(h+1)}$$
$$+\frac{1}{mh(h+1)}\iint\sum_i(\mathbf{x}^{(h+1)}-\mathbf{x}^{(j)}).\frac{\partial}{\partial\mathbf{x}}\sum_i\frac{\partial\phi^{(ih+1)}}{\partial\mathbf{x}^{(i)}}.\frac{\partial F_{h+1}^0}{\partial\boldsymbol{\xi}^{(i)}}d\mathbf{x}^{(h+1)}d\boldsymbol{\xi}^{(h+1)}. \qquad (6\cdot4)$$

As an indication of the general method, this equation will now be discussed for the simplest case $h=1$, when it reduces to

$$-\tfrac{1}{3}\sigma\left(\frac{mv^2}{2kT}-\frac{3}{2}\right)f_1^0 b+\left(\frac{mv^2}{2kT}-\frac{5}{2}\right)\frac{f_1^0}{T}\mathbf{v}.\mathbf{a}+\frac{mf_1^0}{kT}\mathbf{v}.\mathbf{b}.\mathbf{v}=-\frac{1}{m}\iint\frac{\partial\phi}{\partial\mathbf{r}}.\frac{\partial f_2'}{\partial\mathbf{v}^{(1)}}d\mathbf{r}\,d\mathbf{v}^{(2)}. \qquad (6\cdot5)$$

The last term on the right-hand side of (6·4) cancels with the terms involving the pressure on the left-hand side. Apart from a term involving σb, (6·5) is an exact analogue of the corresponding equation in the theory of gases. It is convenient to define the function

$$e=\int f_2'd\mathbf{v}^{(2)}, \qquad (6\cdot6)$$

which must be of the form

$$e=e_0 b+e_{1v}v\mathbf{r}\ \mathbf{a}+e_{1r}r\mathbf{v}.\mathbf{a}+e_{2v}v^2\mathbf{r}.\mathbf{b}.\mathbf{r}+e_{2m}rv\mathbf{r}\ \mathbf{b}.\mathbf{v}+e_{2v}r^2\mathbf{v}.\mathbf{b}.\mathbf{v}. \qquad (6\cdot7)$$

Terms involving b', $\mathbf{a}'$ and $\mathbf{b}'$ may possibly occur, but if they do, (6·5) can clearly provide little information concerning them, because they do not occur on the left-hand side.

A number of conditions must be satisfied by e, for example, since

$$\iint f_2'd\mathbf{x}^{(2)}d\boldsymbol{\xi}^{(2)}=(N-1)f_1',$$
$$-f_1'=\int(e-n_1f_1')d\mathbf{r}. \qquad (6\cdot8)$$

Also
$$n_2'=\int e\,d\mathbf{v}, \qquad (6\cdot9)$$
$$n_2^0\mathbf{u}_2^{(1)\prime}=\int e\mathbf{v}\,d\mathbf{v}. \qquad (6\cdot10)$$

In (6·7), e_0, e_{1r}, e_{1v}, e_{2r}, e_{2m} and e_{2v} are all functions of r, v and z, where

$$rvz = \mathbf{r}\ \mathbf{v}; \tag{6·11}$$

it is convenient to expand them with respect to z in Legendre polynomials, thus:

$$e_k = \sum_l e_k^{(l)} P_l(z), \tag{6·12}$$

where the $e_k^{(l)}$ involve r and v only.

Now substitute from (4·2) and (6·7) into (6·8), making use of integration formulae which will be found in the appendix (§ 7); comparing coefficients of b, $\mathbf{a}$ and $\mathbf{b}$ respectively, it appears that

$$\left.\begin{aligned} \int e_0^{(0)} d\mathbf{r} &= 0, \\ \int (\tfrac{1}{3} e_{1r}^{(1)} r + e_{1v}^{(0)} r - n_1 \varphi_1)\, d\mathbf{r} &= -\varphi_1, \\ \int (\tfrac{1}{5} e_{2r}^{(2)} r^2 + \tfrac{1}{3} e_{2m}^{(1)} r^2 + e_{2v}^{(0)} r^2 - n_1 \varphi_2)\, d\mathbf{r} &= -\varphi_2. \end{aligned}\right\} \tag{6·13}$$

Similarly, from (6·9) and (4·3),

$$\left.\begin{aligned} \int e_0^{(0)} d\mathbf{v} &= 0, \\ \int (e_{1r}^{(0)} + \tfrac{1}{3} e_{1v}^{(1)}) v\, d\mathbf{v} &= 0, \\ \int (e_{2r}^{(0)} + \tfrac{1}{3} e_{2m}^{(1)} + \tfrac{1}{5} e_{2v}^{(2)}) v^2 d\mathbf{v} &= \nu. \end{aligned}\right\} \tag{6·14}$$

To treat (6·10) in the same way, an expression for $N_2^0(\mathbf{U}_2^{(2)\prime} - \mathbf{U}_2^{(1)\prime})$ is required, to combine with (4·4). It must be an odd function of $\mathbf{r}$, and is accordingly of the form

$$N_2^0(\mathbf{U}_2^{(2)\prime} - \mathbf{U}_2^{(1)\prime}) = \upsilon_0 b\mathbf{r} + \upsilon_2 \mathbf{r}.\mathbf{b}.\mathbf{rr} + \bar{\upsilon}_2 r^2 \mathbf{b}\ \mathbf{r}, \tag{6·15}$$

where υ_0, υ_2 and $\bar{\upsilon}_2$ are functions of n, T and r only. It now follows from (6·10) that

$$\left.\begin{aligned} \int \tfrac{1}{3} e_0^{(1)} v\, d\mathbf{v} &= -\tfrac{1}{2} r \upsilon_0, \\ \int (\tfrac{1}{3} e_{1r}^{(1)} + \tfrac{1}{5} e_{1v}^{(2)}) v^2 d\mathbf{v} &= \tfrac{1}{2} r\tau, \\ \int (\tfrac{1}{3} e_{1v}^{(0)} - \tfrac{1}{15} e_{1v}^{(2)}) v^2 d\mathbf{v} &= \tfrac{1}{2} r\tau, \\ \int (\tfrac{1}{3} e_{2r}^{(1)} + \tfrac{1}{5} e_{2m}^{(2)} + \tfrac{1}{7} e_{2v}^{(3)}) v^3 d\mathbf{v} &= -\tfrac{1}{2} r \upsilon_2, \\ \int (\tfrac{1}{3} e_{2m}^{(0)} + \tfrac{2}{15} e_{2v}^{(1)} - \tfrac{1}{15} e_{2m}^{(2)} - \tfrac{2}{35} e_{2v}^{(3)}) v^3 d\mathbf{v} &= -\tfrac{1}{2} r \upsilon_2 \end{aligned}\right\} \tag{6·16}$$

According to (3·30),

$$\upsilon_0 = r^{-3} \int_0^r \left(n \frac{\partial N_2^0}{\partial n} - N_2^0 + \frac{\sigma + 2}{3} T \frac{\partial N_2^0}{\partial T} \right) r^2 dr, \tag{6·17}$$

and

$$\frac{\partial}{\partial \mathbf{r}} . (\upsilon_2 \mathbf{r}.\mathbf{b}.\mathbf{rr} + \bar{\upsilon}_2 r^2 \mathbf{b}\ \mathbf{r}) = 0, \tag{6·18}$$

so that

$$\frac{d}{dr}(r^5 \upsilon_2) + r^3 \frac{d}{dr}(r^2 \bar{\upsilon}_2) = 0 \tag{6·19}$$

Now the integral equation separates into the following three:

$$\frac{\partial}{\partial \mathbf{v}} \cdot \int \frac{\partial \phi}{\partial \mathbf{r}} e_0 d\mathbf{r} = \frac{m\sigma}{3}\left(\frac{mv^2}{2kT} - \frac{3}{2}\right) f_1^0, \tag{6·20}$$

$$\frac{\partial}{\partial \mathbf{v}} \cdot \int \frac{\partial \phi}{\partial \mathbf{r}} (e_{1r} v\mathbf{r} + e_{1v} r\mathbf{v})\, d\mathbf{r} = -\frac{m f_1^0}{T}\left(\frac{mv^2}{2kT} - \frac{5}{2}\right)\mathbf{v}, \tag{6·21}$$

$$\frac{\partial}{\partial \mathbf{v}} \cdot \int \frac{\partial \phi}{\partial \mathbf{r}} \{e_{2r} v^2(\mathbf{rr} - \tfrac{1}{3}r^2 1) + e_{2m} rv(\tfrac{1}{2}\mathbf{rv} + \tfrac{1}{2}\mathbf{vr} - \tfrac{1}{3}\mathbf{r}.\mathbf{v}\, 1) + e_{2v} r^2(\mathbf{vv} - \tfrac{1}{3}v^2 1)\}\, d\mathbf{r}$$
$$= -\frac{m^2 f_1^0}{kT}(\mathbf{vv} - \tfrac{1}{3}v^2 1). \tag{6·22}$$

Since $\frac{\partial}{\partial \mathbf{v}} \cdot (v^{-5}\mathbf{v}) = 0$, the integral of (6·20) is

$$\int \phi'(r)\, e_0^{(1)} d\mathbf{r} = \frac{m\sigma}{v^3} \int_\infty^v v^3 \left(\frac{mv^2}{2kT} - \frac{3}{2}\right) f_1^0 dv. \tag{6·23}$$

Also, since $\frac{\partial}{\partial \mathbf{v}} \cdot \{v^{-5}(\mathbf{vv} - \frac{1}{3}v^2 1)\} = 0$, it follows from (6 21) that

$$\int \phi'(r)\,(\tfrac{1}{3}e_{1v}^{(1)} + \tfrac{1}{5}e_{1r}^{(2)})\, r\, d\mathbf{r} = -\frac{3m}{2Tv^4} \int_\infty^v v^4 \left(\frac{mv^2}{2kT} - \frac{5}{2}\right) f_1^0 dv. \tag{6·24}$$

Finally, since $\sum_\alpha \frac{\partial}{\partial v_\alpha}[v^{-7}\{v_\alpha v_\beta v_\gamma - \frac{1}{5}v^2(v_\alpha \delta_{\beta\gamma} + v_\beta \delta_{\gamma\alpha} + v_\gamma \delta_{\alpha\beta})\}] = 0$,

one obtains from (6·22)

$$\int \phi'(r)\,(\tfrac{1}{3}e_{1v}^{(1)} + \tfrac{1}{5}e_{2m}^{(2)} + \tfrac{1}{7}e_{2r}^{(2)})\, r^2 d\mathbf{r} = -\frac{5m^2}{3kTv^6} \int_\infty^v v^6 f_1^0 dv. \tag{6·25}$$

These relations do not suffice to determine the function e completely without other assumptions; however, if one assumes, as seems reasonable, e_0, e_1 and e_2 each involve the velocity as a simple factor (self-consistent field), it follows from (6 23) to (6 25) that

$$\left.\begin{aligned} e_0 &\sim \frac{1}{v^3}\int_\infty^v v^3 \left(\frac{mv^2}{2kT} - \frac{3}{2}\right) f_1^0 dv, \\ e_1 &\sim \frac{1}{v^4}\int_\infty^v v^4 \left(\frac{mv^2}{2kT} - \frac{5}{2}\right) f_1^0 dv, \\ e_2 &\sim \frac{1}{v^5}\int_\infty^v v^5 f_1^0 dv. \end{aligned}\right\} \tag{6·26}$$

Then the velocity dependence of e is determined by (6·4) with $h = 1$ in terms of incomplete gamma-functions; to obtain the dependence on r, it will be necessary to proceed to the equation $h = 2$, using the 'Kirkwood' approximation [(4·6) of part I] for f_3 in terms of f_2. When e has been determined, ν and τ are given by (6·14) and (6·16) respectively, and the coefficients of viscosity and thermal conduction may be accurately calculated.

7. Appendix

Here are derived some formulae of integration which are required in § 6. The following properties of Legendre polynomials are involved

$$\frac{1}{2}\int_{-1}^{1} P_m(z)\,P_n(z)\,dz = \frac{\delta_{mn}}{2m+1}, \tag{7·1}$$

and

$$\left.\begin{aligned} P_0(z) &= 1, \\ P_1(z) &= z, \\ P_2(z) &= \tfrac{1}{2}(3z^2-1), \\ P_3(z) &= \tfrac{1}{2}z(5z^2-3). \end{aligned}\right\} \tag{7·2}$$

It follows immediately that

$$\begin{aligned} \int e_k\,d\mathbf{r} &= \int_0^\infty \frac{1}{2}\int_{-1}^{1} P_0(z)\sum_l e_k^{(l)}P_l(z)\,dz\,.\,4\pi r^2\,dr \\ &= \int e_k^{(0)}\,d\mathbf{r}, \end{aligned} \tag{7·3}$$

and

$$\begin{aligned} \int e_k\,\mathbf{r}\,d\mathbf{r} &= v^{-1}\,\mathbf{v}\int e_k r z\,d\mathbf{r} \\ &= v^{-1}\,\mathbf{v}\int_0^\infty \frac{1}{2}\int_{-1}^{1} P_1(z)\sum_l e_k^{(l)}P_l(z)\,dz\,r\,.\,4\pi r^2\,dr \\ &= \tfrac{1}{3}v^{-1}\,\mathbf{v}\int e_k^{(1)} r\,d\mathbf{r}. \end{aligned} \tag{7·4}$$

Next, it is seen that

$$\begin{aligned} \int e_k\,\mathbf{r}\mathbf{r}\,d\mathbf{r} &= \int e_k(\mathbf{r}\mathbf{r}-\tfrac{1}{3}r^2\,1)\,d\mathbf{r}+\frac{1}{3}\int e_k r^2\,d\mathbf{r}\,1 \\ &= v^{-2}(\mathbf{v}\mathbf{v}-\tfrac{1}{3}v^2\,1)\int e_k\tfrac{1}{2}(3z^2-1)\,r^2\,d\mathbf{r}+\frac{1}{3}\int e_k r^2\,d\mathbf{r}\,1 \\ &= \tfrac{1}{5}v^{-2}(\mathbf{v}\mathbf{v}-\tfrac{1}{3}v^2\,1)\int e_k^{(2)} r^2\,d\mathbf{r}+\tfrac{1}{3}\,1\int e_k^{(0)} r^2\,d\mathbf{r}. \end{aligned} \tag{7·5}$$

In a similar way,

$$\begin{aligned} \int e_k r_\alpha r_\beta r_\gamma\,d\mathbf{r} &= \int e_k\{r_\alpha r_\beta r_\gamma-\tfrac{1}{5}r^2(r_\alpha\delta_{\beta\gamma}+r_\beta\delta_{\gamma\alpha}+r_\gamma\delta_{\alpha\beta})\}\,d\mathbf{r} \\ &\quad+\frac{1}{5}\int e_k r^2(r_\alpha\delta_{\beta\gamma}+r_\beta\delta_{\gamma\alpha}+r_\gamma\delta_{\alpha\beta})\,d\mathbf{r} \\ &= v^{-3}\{v_\alpha v_\beta v_\gamma-\tfrac{1}{5}v^2(v_\alpha\delta_{\beta\gamma}+v_\beta\delta_{\gamma\alpha}+v_\gamma\delta_{\alpha\beta})\}\int e_k\tfrac{1}{2}(5z^3-3z)\,r^3\,d\mathbf{r} \\ &\quad+\tfrac{1}{5}v^{-1}(v_\alpha\delta_{\beta\gamma}+v_\beta\delta_{\gamma\alpha}+v_\gamma\delta_{\alpha\beta})\int e_k z r^3\,d\mathbf{r} \\ &= \tfrac{1}{7}v^{-3}\{v_\alpha v_\beta v_\gamma-\tfrac{1}{5}v^2(v_\alpha\delta_{\beta\gamma}+v_\beta\delta_{\gamma\alpha}+v_\gamma\delta_{\alpha\beta})\}\int e_k^{(3)} r^3\,d\mathbf{r} \\ &\quad+\tfrac{1}{15}v^{-1}(v_\alpha\delta_{\beta\gamma}+v_\beta\delta_{\gamma\alpha}+v_\gamma\delta_{\alpha\beta})\int e_k^{(1)} r^3\,d\mathbf{r}. \end{aligned} \tag{7·6}$$

It is obvious that the formulae of this section remain valid when $\mathbf{r}$ and $\mathbf{v}$ are interchanged.

References

Andrade, E. N. da C. 1934 *Phil. Mag.* ser. 7, **17**, 497, 698.
Born, M. & Green, H. S. 1946 *Proc. Roy. Soc.* A, **188**, 10.
Chapman, S. 1912 *Phil. Trans.* A, **211**, 433.
Chapman, S. 1916 *Phil. Trans.* A, **216**, 279.
Chapman, S. & Cowling, T. G. 1939 *The mathematical theory of non-uniform gases.* Camb. Univ. Press.
Enskog, D. 1911 *Phys. Z.* **12**, 56, 533.
Enskog, D. 1917 Uppsala Diss
Ewell, R. H. & Eyring, H. 1937 *J. Chem. Phys.* **5**, 726.
Fürth, R. 1941 *Proc. Camb. Phil. Soc.* **37**, 281.
Green, H. S. 1947 *Proc. Roy. Soc.* A, **189**, 103.
de Guzman Carrancio J. 1913 *An. Soc. Esp. Fis. Quim.* **11**, 353.
Hilbert, D. 1912 *Math. Ann* **72**, 562.
Kahn, B. 1938 Utrecht Diss.
Lennard-Jones, J. E. 1924 *Proc. Roy. Soc.* A, **106**, 463.

A dynamical model of a crystal structure

By Sir Lawrence Bragg, F.R.S. and J. F. Nye

Cavendish Laboratory, University of Cambridge

(*Received* 9 *January* 1947—*Read* 19 *June* 1947)

[Plates 8 to 21]

The crystal structure of a metal is represented by an assemblage of bubbles, a millimetre or less in diameter, floating on the surface of a soap solution. The bubbles are blown from a fine pipette beneath the surface with a constant air pressure, and are remarkably uniform in size. They are held together by surface tension, either in a single layer on the surface or in a three-dimensional mass. An assemblage may contain hundreds of thousands of bubbles and persists for an hour or more. The assemblages show structures which have been supposed to exist in metals, and simulate effects which have been observed, such as grain boundaries, dislocations and other types of fault, slip, recrystallization, annealing, and strains due to 'foreign' atoms.

1. The bubble model

Models of crystal structure have been described from time to time in which the atoms are represented by small floating or suspended magnets, or by circular disks floating on a water surface and held together by the forces of capillary attraction. These models have certain disadvantages; for instance, in the case of floating objects in contact, frictional forces impede their free relative movement. A more serious disadvantage is that the number of components is limited, for a large number of components is required in order to approach the state of affairs in a real crystal

The present paper describes the behaviour of a model in which the atoms are represented by small bubbles from 2·0 to 0·1 mm. in diameter floating on the surface of a soap solution. These small bubbles are sufficiently persistent for experiments lasting an hour or more, they slide past each other without friction, and they can be produced in large numbers Some of the illustrations in this paper were taken from assemblages of bubbles numbering 100,000 or more. The model most nearly represents the behaviour of a metal structure, because the bubbles are of one type only and are held together by a general capillary attraction, which represents the binding force of the free electrons in the metal. A brief description of the model has been given in the *Journal of Scientific Instruments* (Bragg 1942*b*)

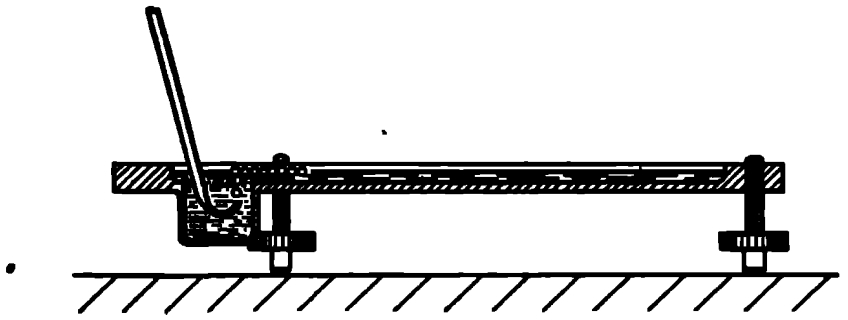


FIGURE 1. Apparatus for producing rafts of bubbles

2. METHOD OF FORMATION

The bubbles are blown from a fine orifice, beneath the surface of a soap solution. We have had the best results with a solution the formula of which was given to us by Mr Green of the Royal Institution. 15·2 c c of oleic acid (pure redistilled) is well shaken in 50 c c of distilled water. This is mixed thoroughly with 73 c c. of 10 % solution of tri-ethanolamine and the mixture made up to 200 c.c. To this is added 164 c.c of pure glycerine. It is left to stand and the clear liquid is drawn off from below. In some experiments this was diluted in three times its volume of water to reduce viscosity. The orifice of the jet is about 5 mm below the surface. A constant air pressure of 50 to 200 cm. of water is supplied by means of two Winchester flasks. Normally the bubbles are remarkably uniform in size. Occasionally they issue in an irregular manner, but this can be corrected by a change of jet or of pressure. Unwanted bubbles can easily be destroyed by playing a small flame over the surface. Figure 1 shows the apparatus. We have found it of advantage to blacken the bottom of the vessel, because details of structure, such as grain boundaries and dislocations, then show up more clearly

Figure 2, plate 8, shows a portion of a 'raft' or two-dimensional crystal of bubbles. Its regularity can be judged by looking at the figure in a glancing direction. The size of the bubbles varies with the aperture, but does not appear to vary to any marked degree with the pressure or the depth of the orifice beneath the surface.

The main effect of increasing the pressure is to increase the rate of issue of the bubbles. As an example, a thick-walled jet of 49μ bore with a pressure of 100 cm. produced bubbles of 1·2 mm in diameter. A thin-walled jet of 27μ diameter and a pressure of 180 cm produced bubbles of 0·6 mm diameter. It is convenient to refer to bubbles of 2·0 to 1·0 mm. diameter as 'large' bubbles, those from 0·8 to 0·6 mm. diameter as 'medium' bubbles, and those from 0·3 to 0·1 mm. diameter as 'small' bubbles, since their behaviour varies with their size.

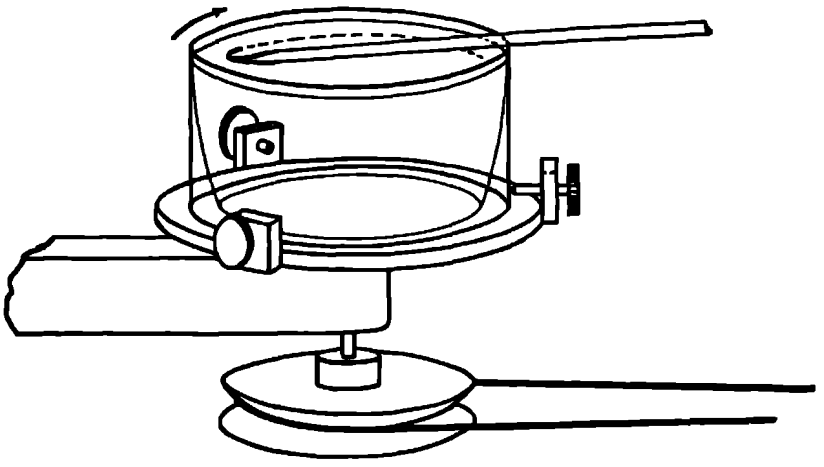


FIGURE 3. Apparatus for producing bubbles of small size.

With this apparatus we have not found it possible to reduce the size of the jet and so produce bubbles of smaller diameter than 0·6 mm. As it was desired to experiment with very small bubbles, we had recourse to placing the soap solution in a rotating vessel and introducing a fine jet as nearly as possible parallel to a stream line. The bubbles are swept away as they form, and under steady conditions are reasonably uniform. They issue at a rate of one thousand or more per second, giving a high-pitched note. The soap solution mounts up in a steep wall around the perimeter of the vessel while it is rotating, but carries back most of the bubbles with it when rotation ceases. With this device, illustrated in figure 3, bubbles down to 0·12 mm. in diameter can be obtained. As an example, an orifice 38μ across in a thin-walled jet, with a pressure of 190 cm of water, and a speed of the fluid of 180 cm /sec past the orifice, produced bubbles of 0 14 mm diameter. In this case a dish of diameter 9·5 cm. and speed of 6 rev /sec was used. Figure 4, plate 8, is an enlarged picture of these 'small' bubbles and shows their degree of regularity, the pattern is not as perfect with a rotating as with a stationary vessel, the rows being seen to be slightly irregular when viewed in a glancing direction.

These two-dimensional crystals show structures which have been supposed to exist in metals, and simulate effects which have been observed, such as grain boundaries, dislocations and other types of fault, slip, recrystallization, annealing, and strains due to 'foreign' atoms.

3 Grain boundaries

Figures 5*a*, 5*b* and 5*c*, plates 9 and 10, show typical grain boundaries for bubbles of 1·87, 0·76 and 0·30 mm. diameter respectively. The width of the disturbed area at the boundary, where the bubbles have an irregular distribution, is in general greater the smaller the bubbles. In figure 5*a*, which shows portions of several adjacent grains, bubbles at a boundary between two grains adhere definitely to one crystalline arrangement or the other. In figure 5*c* there is a marked 'Beilby layer' between the two grains. The small bubbles, as will be seen, have a greater rigidity than the large ones, and this appears to give rise to more irregularity at the interface.

Separate grains show up distinctly when photographs of polycrystalline rafts such as figures 5*a* to 5*c*, plates 9 and 10, and figures 12*a* to 12*e*, plates 14 to 16, are viewed obliquely. With suitable lighting, the floating raft of bubbles itself when viewed obliquely resembles a polished and etched metal in a remarkable way.

It often happens that some 'impurity atoms', or bubbles which are markedly larger or smaller than the average, are found in a polycrystalline raft, and when this is so a large proportion of them are situated at the grain boundaries It would be incorrect to say that the irregular bubbles make their way to the boundaries, it is a defect of the model that no diffusion of bubbles through the structure can take place, mutual adjustments of neighbours alone being possible. It appears that the boundaries tend to readjust themselves by the growth of one crystal at the expense of another till they pass through the irregular atoms.

4. Dislocations

When a single crystal or polycrystalline raft is compressed, extended, or otherwise deformed it exhibits a behaviour very similar to that which has been pictured for metals subjected to strain. Up to a certain limit the model is within its elastic range. Beyond that point it yields by slip along one of the three equally inclined directions of closely packed rows. Slip takes place by the bubbles in one row moving forward over those in the next row by an amount equal to the distance between neighbours It is very interesting to watch this process taking place. The movement is not simultaneous along the whole row but begins at one end with the appearance of a 'dislocation', where there is locally one more bubble in the rows on one side of the slip line as compared with those on the other. This dislocation then runs along the slip line from one side of the crystal to the other, the final result being a slip by one 'inter-atomic' distance Such a process has been invoked by Orowan, by Polanyi and by Taylor to explain the small forces required to produce plastic gliding in metal structures. The theory put forward by Taylor (1934) to explain the mechanism of plastic deformation of crystals considers the mutual action and equilibrium of such dislocations. The bubbles afford a very striking picture of what has been supposed to take place in the metal. Sometimes the dislocations run along quite slowly, taking a matter of seconds to cross a crystal; stationary dislocations also are to be seen in crystals which are not homogeneously

strained. They appear as short black lines, and can be seen in the series of photographs, figures 12*a* to 12*e*, plates 14 to 16. When a polycrystalline raft is compressed, these dark lines are seen to be dashing about in all directions across the crystals

Figures 6*a*, 6*b* and 6*c*, plates 10 and 11, show examples of dislocations. In figure 6*a*, where the diameter of the bubbles is 1·9 mm , the dislocation is very local, extending over about six bubbles. In figure 6*b* (diameter 0·76 mm.) it extends over twelve bubbles, and in figure 6*c* (diameter 0·30 mm) its influence can be traced for a length of about fifty bubbles. The greater rigidity of the small bubbles leads to longer dislocations. The study of any mass of bubbles shows, however, that there is not a standard length of dislocation for each size. The length depends upon the nature of the strain in the crystal. A boundary between two crystals with corresponding axes at approximately 30° (the maximum angle which can occur) may be regarded as a series of dislocations in alternate rows, and in this case the dislocations are very short. As the angle between the neighbouring crystals decreases, the dislocations occur at wider intervals and at the same time become longer, till one finally has single dislocations in a large body of perfect structure as shown in figures 6*a*, 6*b* and 6*c*.

Figure 7, plate 11, shows three parallel dislocations If we call them positive and negative (following Taylor) they are positive, negative, positive, reading from left to right. The strip between the last two has three bubbles in excess, as can be seen by looking along the rows in a horizontal direction. Figure 8, plate 12, shows a dislocation projecting from a grain boundary, an effect often observed.

Figure 9, plate 12, shows a place where two bubbles take the place of one. This may be regarded as a limiting case of positive and negative dislocations on neighbouring rows, with the compressive sides of the dislocations facing each other. The contrary case would lead to a hole in the structure, one bubble being missing at the point where the dislocations met.

5. Other types of fault

Figure 10, plate 12, shows a narrow strip between two crystals of parallel orientation, the strip being crossed by a number of fault lines where the bubbles are not in close packing It is in such places as these that recrystallization may be expected The boundaries approach and the strip is absorbed into a wider area of perfect crystal

Figures 11*a* to 11*g*, plates 13 and 14 are examples of arrangements which frequently appear in places where there is a local deficiency of bubbles. While a dislocation is seen as a dark stripe in a general view, these structures show up in the shape of the letter V or as triangles. A typical V structure is seen in figure 11*a*. When the model is being distorted, a V structure is formed by two dislocations meeting at an inclination of 60°; it is destroyed by the dislocations continuing along their paths. Figure 11*b* shows a small triangle, which also embodies a dislocation, for it will be noticed that the rows below the fault have one more bubble than those below. If a mild amount

of 'thermal movement' is imposed by gentle agitation of one side of the crystal, such faulty places disappear and a perfect structure is formed.

Here and there in the crystals there is a blank space where a bubble is missing, showing as a black dot in a general view. Examples occur in figure 11*g*. Such a gap cannot be closed by a local readjustment, since filling the hole causes another to appear. Such holes both appear and disappear when the crystal is 'cold-worked'.

These structures in the model suggest that similar local faults may exist in an actual metal. They may play a part in processes such as diffusion or the order-disorder change by reducing energy barriers in their neighbourhood, and act as nuclei for crystallization in an allotropic change

6. Recrystallization and annealing

Figures 12*a* to 12*e*, plates 14 to 16, show the same raft of bubbles at successive times. A raft covering the surface of the solution was given a vigorous stirring with a glass rake, and then left to adjust itself. Figure 12*a* shows its aspect about 1 sec. after stirring has ceased. The raft is broken into a number of small 'crystallites', these are in a high state of non-homogeneous strain as is shown by the numerous dislocations and other faults. The following photograph (figure 12*b*) shows the same raft 32 sec later. The small grains have coalesced to form larger grains, and much of the strain has disappeared in the process. Recrystallization takes place right through the series, the last three photographs of which show the appearance of the raft 2, 14 and 25 min after the initial stirring. It is not possible to follow the rearrangement for much longer times, because the bubbles shrink after long standing, apparently due to the diffusion of air through their walls, and they also become thin and tend to burst No agitation was given to the model during this process. An ever slower process of rearrangement goes on, the movement of the bubbles in one part of the raft setting up strains which activate a rearrangement in a neighbouring part, and that in its turn still another

A number of interesting points are to be seen in this series. Note the three small grains at the points indicated by the co-ordinates *AA*, *BB*, *CC*. *A* persists, though changed in form, throughout the whole series. *B* is still present after 14 min., but has disappeared in 25 min , leaving behind it four dislocations marking internal strain in the grain. Grain *C* shrinks and finally disappears in figure 12*d*, leaving a hole and a V which has disappeared in figure 12*e* At the same time the ill-defined boundary in figure 12*d* at *DD* has become a definite one in figure 12*e* Note also the straightening out of the grain boundary in the neighbourhood of *EE* in figures 12*b* to 12*e*. Dislocations of various lengths can be seen, marking all stages between a slight warping of the structure and a definite boundary. Holes where bubbles are missing show up as black dots. Some of these holes are formed or filled up by movements of dislocations, but others represent places where a bubble has burst. Many examples of V's and some of triangles can be seen Other interesting points will be apparent from a study of this series of photographs.

Figures 13*a*, 13*b* and 13*c*, plate 17, show a portion of a raft 1 sec., 4 sec. and 4 min. after the stirring process, and is interesting as showing two successive stages in the relaxation towards a more perfect arrangement. The changes show up well when one looks in a glancing direction across the page The arrangement is very broken in figure 13*a* In figure 13*b* the bubbles have grouped themselves in rows, but the curvature of these rows indicates a high degree of internal strain. In figure 13*c* this strain has been relieved by the formation of a new boundary at *A*–*A*, the rows on either side now being straight. It would appear that the energy of this strained crystal is greater than that of the intercrystalline boundary. We are indebted to Messrs Kodak for the photographs of figure 13, which were taken when the cinematograph film referred to below was produced.

7. Effect of impurity atom

Figure 14, plate 18, shows the widespread effect of a bubble which is of the wrong size. If this figure is compared with the perfect rafts shown in figures 2 and 4, plate 8, it will be seen that three bubbles, one larger and two smaller than normal, disturb the regularity of the rows over the whole of the figure. As has been mentioned above, bubbles of the wrong size are generally found in the grain boundaries, where holes of irregular size occur which can accommodate them.

8. Mechanical properties of the two-dimensional model

The mechanical properties of a two-dimensional perfect raft have been described in the paper referred to above (Bragg 1942*b*) The raft lies between two parallel springs dipping horizontally in the surface of the soap solution. The pitch of the springs is adjusted to fit the spacing of the rows of bubbles, which then adhere firmly to them. One spring can be translated parallel to itself by a micrometer screw, and the other is supported by two thin vertical glass fibres. The shearing stress can be measured by noting the deflexion of the glass fibres When subjected to a shearing strain, the raft obeys Hooke's law of elasticity up to the point where the elastic limit is reached It then slips along some intermediate row by an amount equal to the width of one bubble The elastic shear and slip can be repeated several times The elastic limit is approximately reached when one side of the raft has been sheared by an amount equal to a bubble width past the other side. This feature supports the basic assumption made by one of us in the calculation of the elastic limit of a metal (Bragg 1942*a*), in which it is supposed that each crystallite in a cold-worked metal only yields when the strain in it has reached such a value that energy is released by the slip

A calculation has been made by M. M Nicolson of the forces between the bubbles, and will be published shortly. It shows two interesting points. The curve for the variation of potential energy with distance between centres is very similar to those which have been plotted for atoms. It has a minimum for a distance between centres slightly less than a free bubble diameter, and rises sharply for smaller distances. Further, the rise is extremely sharp for bubbles of 0·1 mm. diameter but

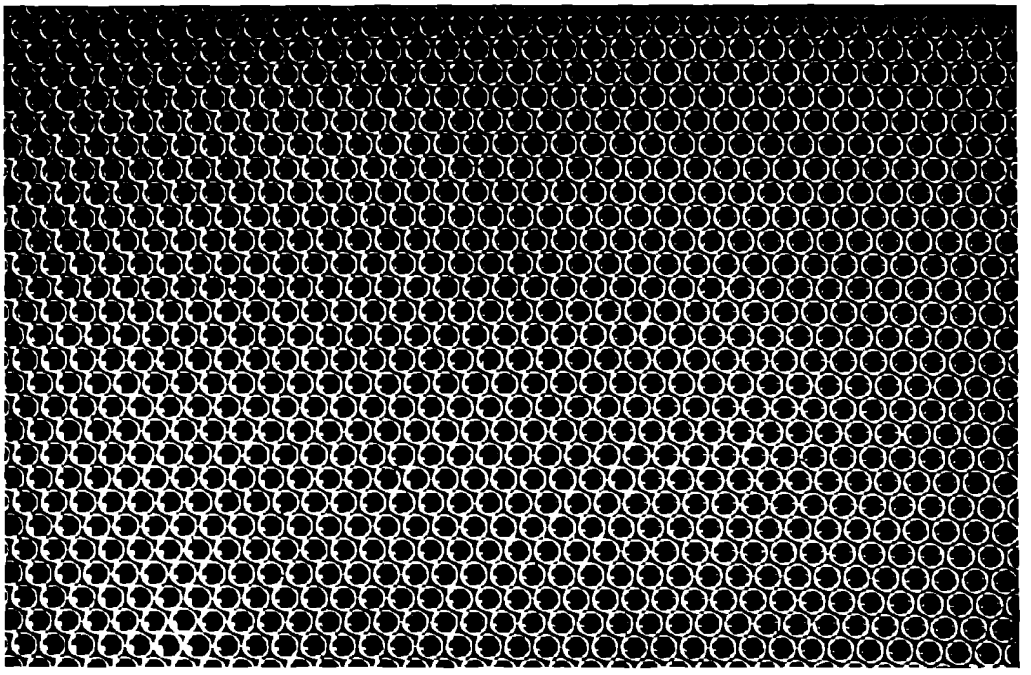


FIGURE 2 Perfect crystalline raft of bubbles Diameter 1.41 mm

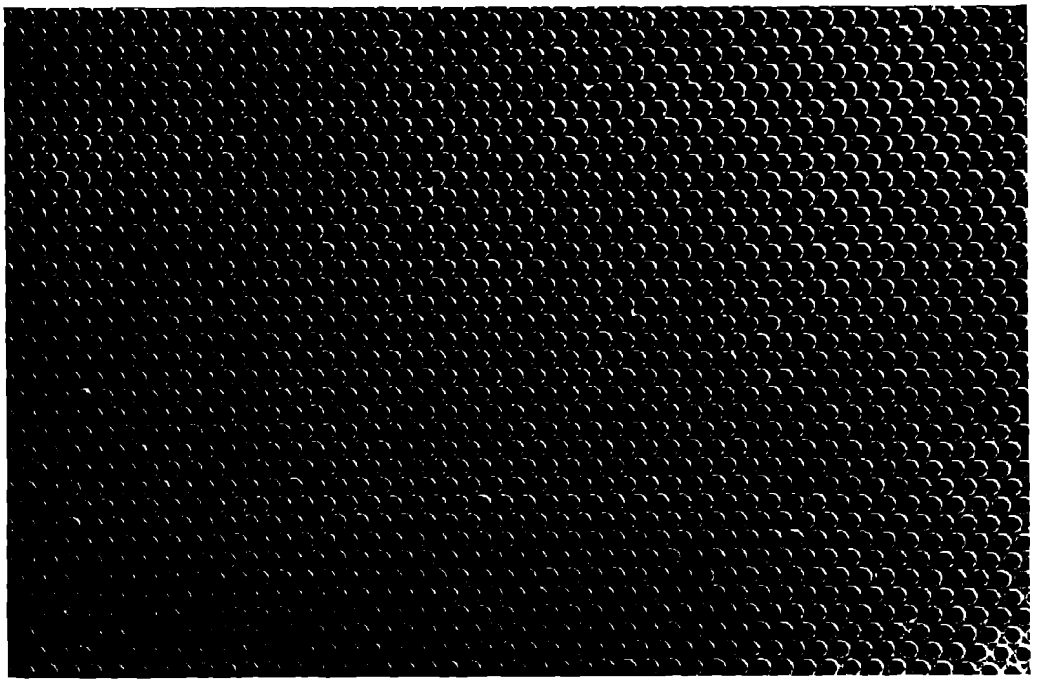


FIGURE 4 Perfect crystalline raft of bubbles Diameter 0.30 mm

(*Facing p.* 180)

Grain boundaries

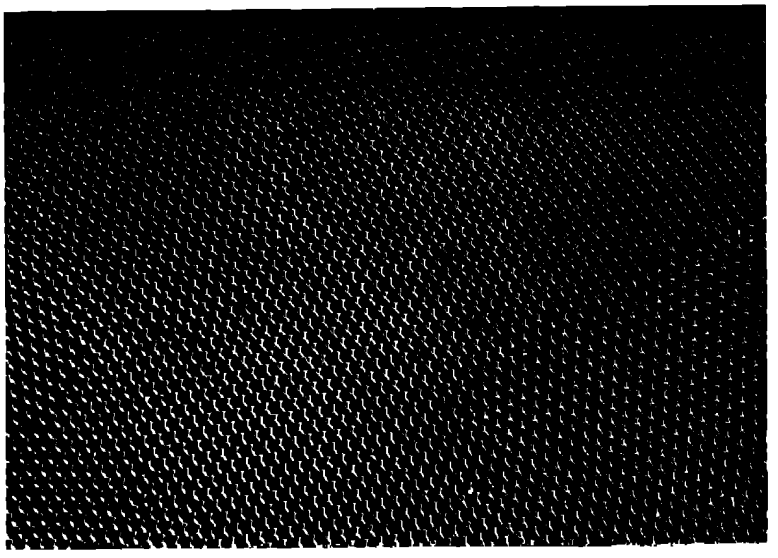


FIGURE 5*a* Diameter 1·87 mm

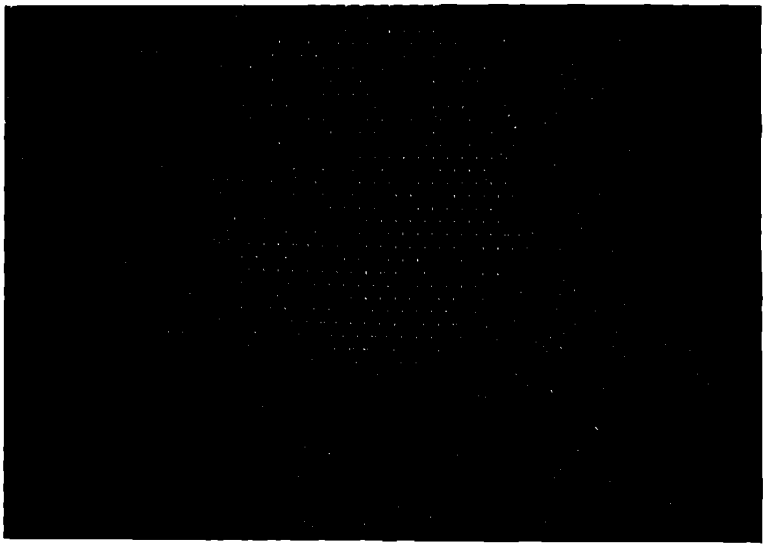


FIGURE 5*b* Diameter 0·76 mm

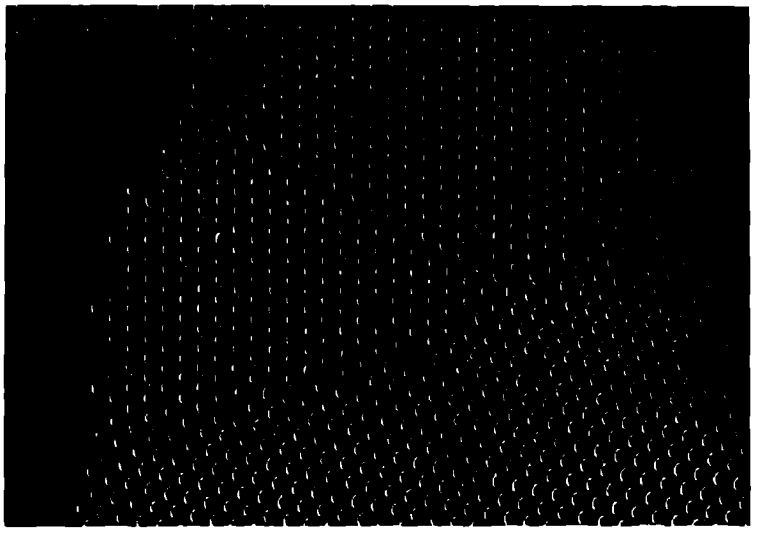


FIGURE 5*a*. A grain boundary. Diameter 0·10 mm

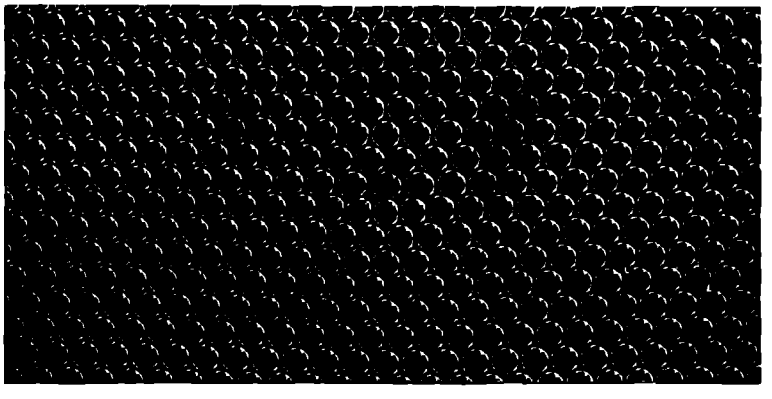


FIGURE 6*a*. A dislocation. Diameter 1·9 mm

Dislocations

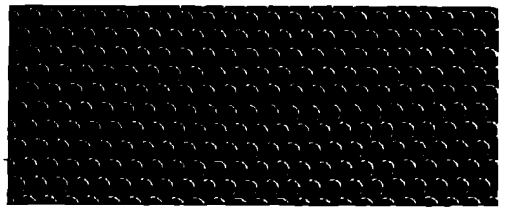


FIGURE 6*b* Diameter 0·76 mm

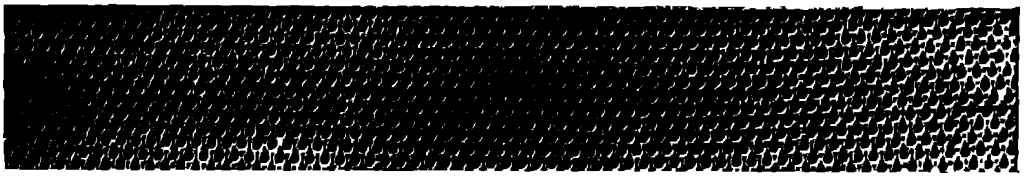


FIGURE 6*c* Diameter 0·30 mm

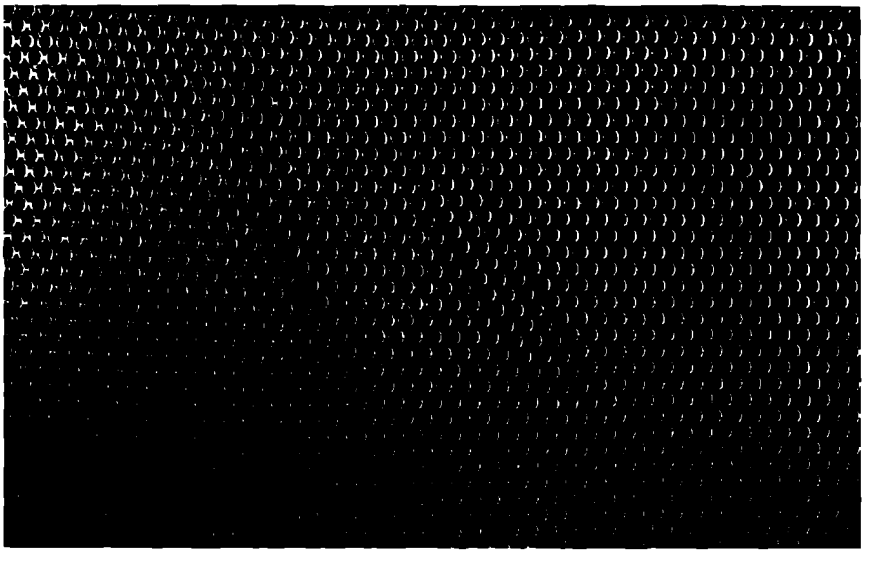


FIGURE 7 Parallel dislocations Diameter 0·76 mm

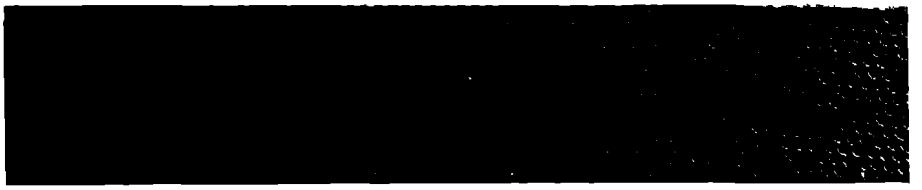


FIGURE 8 Dislocation projecting from a grain boundary Diameter 0 30 mm.

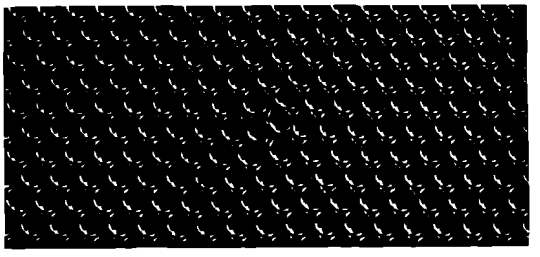


FIGURE 9 Dislocations in adjacent rows Diameter 1 9 mm

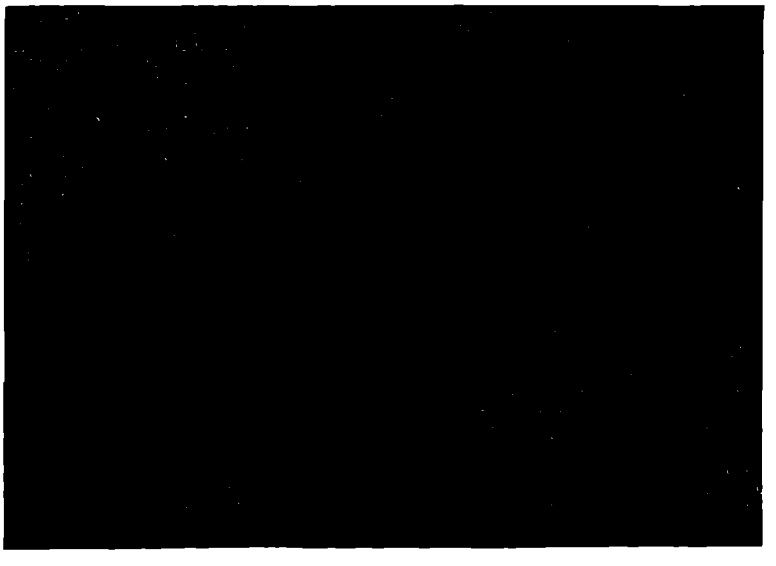
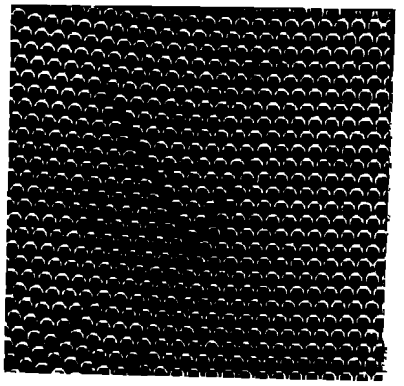
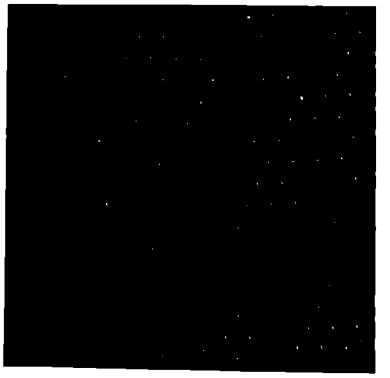


FIGURE 10 Series of fault lines between two areas of parallel orientation Diameter 0 30 mm



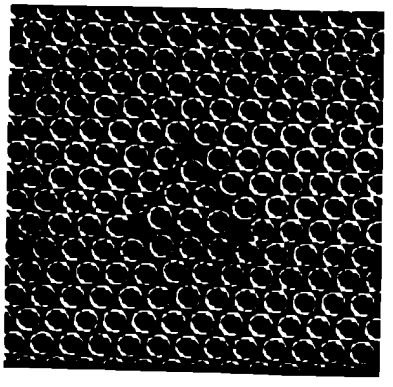
Diameter 0·68 mm

a



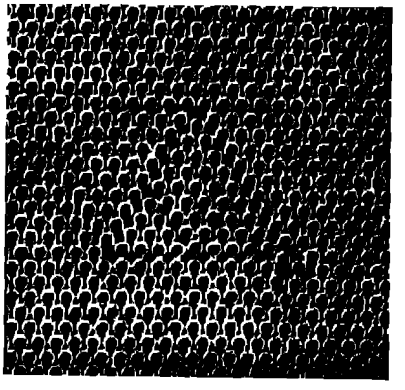
Diameter 0·68 mm

b



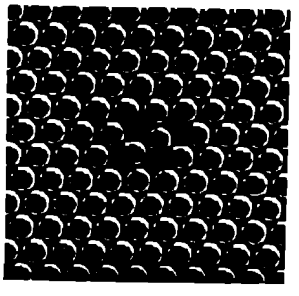
Diameter 0·6 mm

c



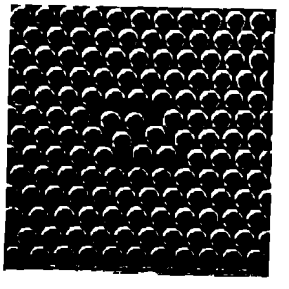
Diameter 0·30 mm

d



Diameter 0·6 mm

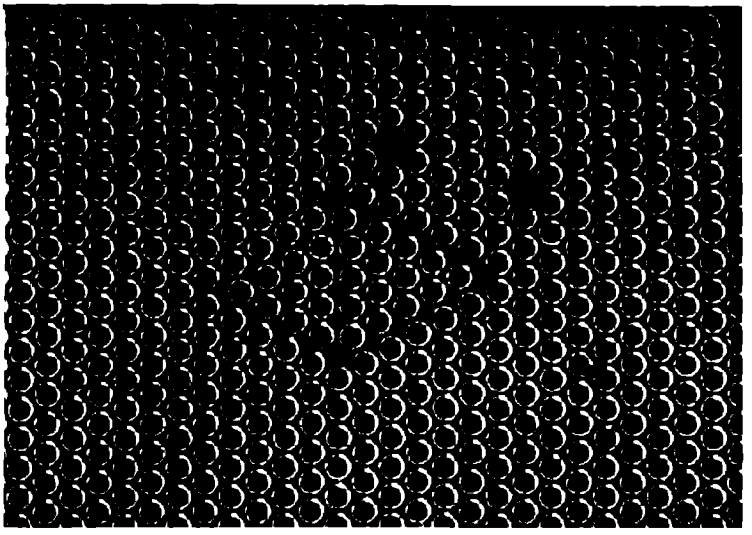
e



Diameter 0·6 mm

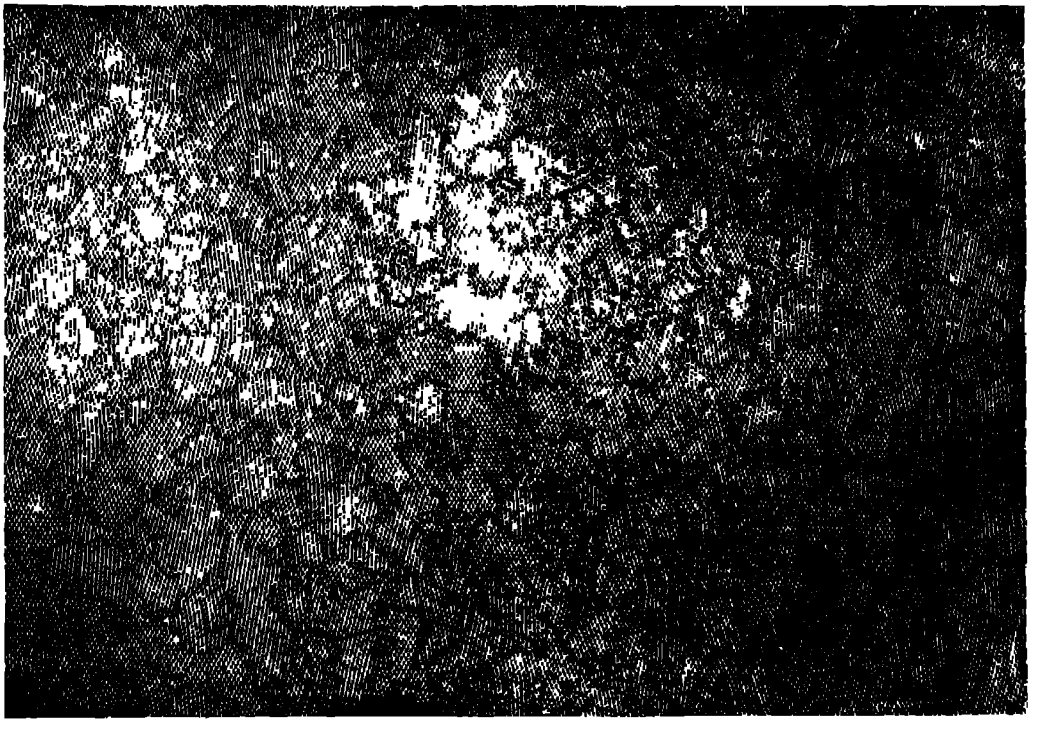
f

FIGURE 11 Types of fault



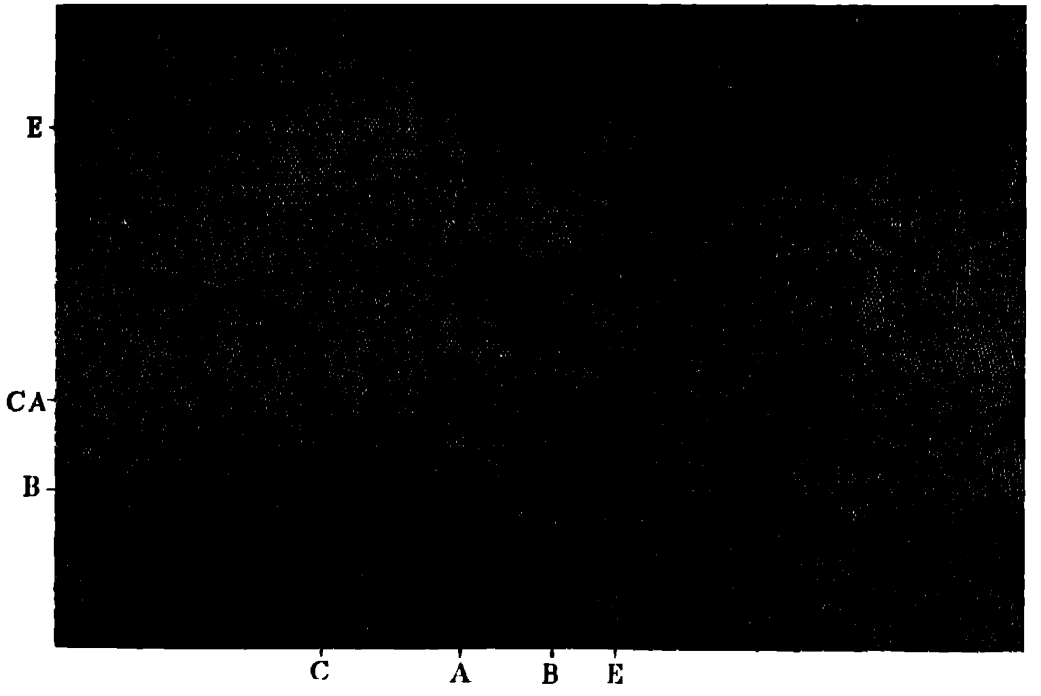
q

FIGURE 11. Types of fault. Diameter 0·68 mm.

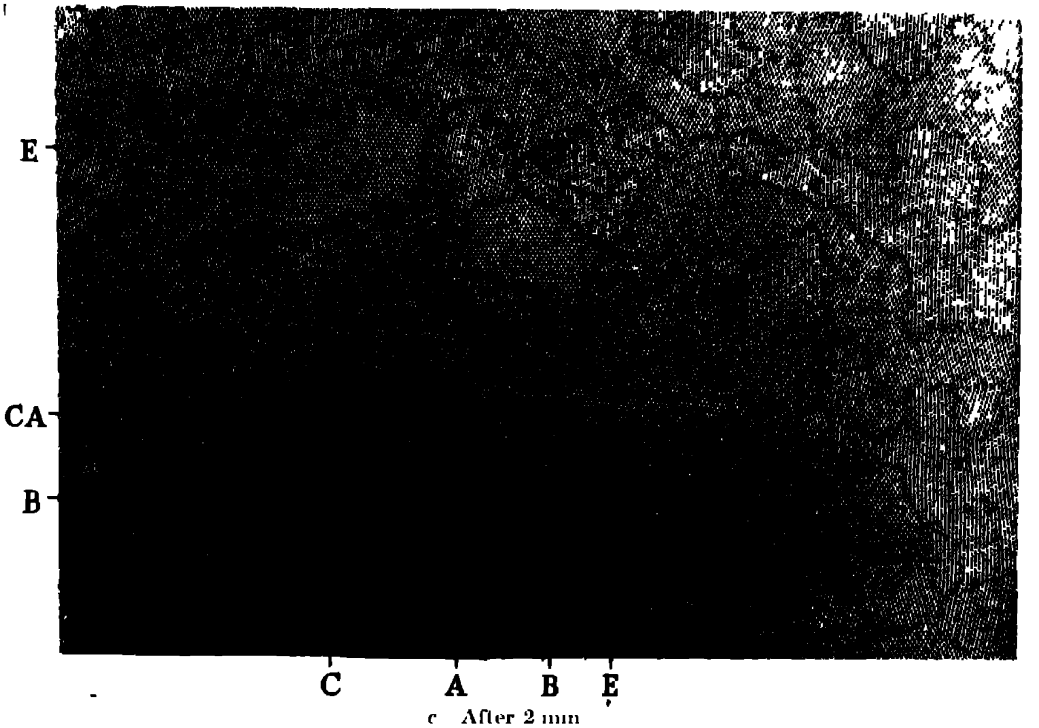


a Immediately after stirring.

FIGURE 12. Recrystallization. Diameter 0·60 mm.

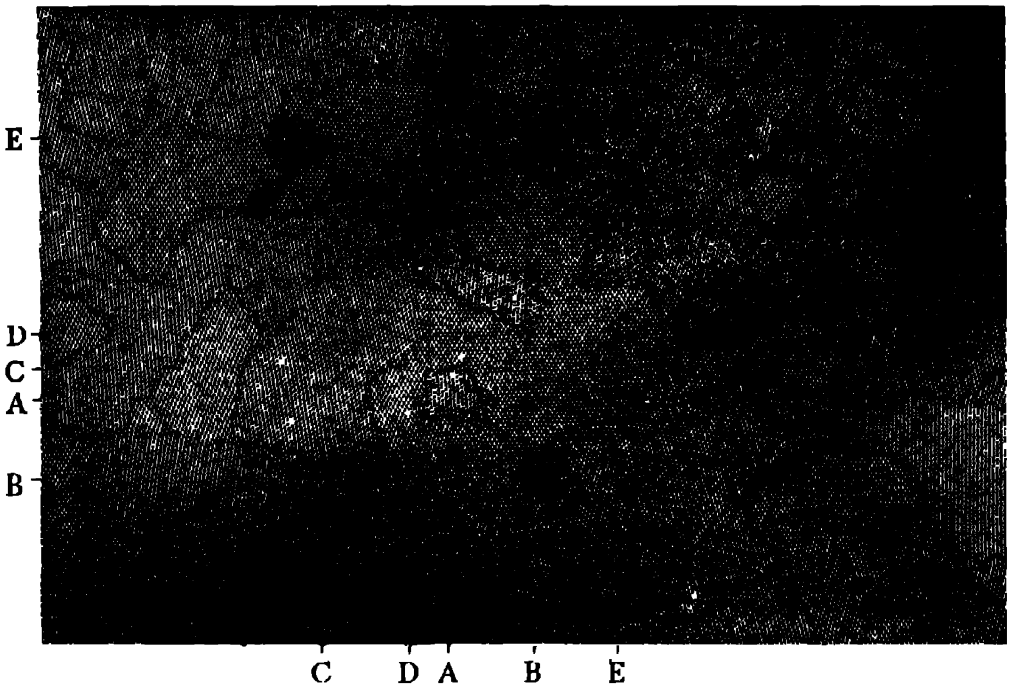


b After 33 sec

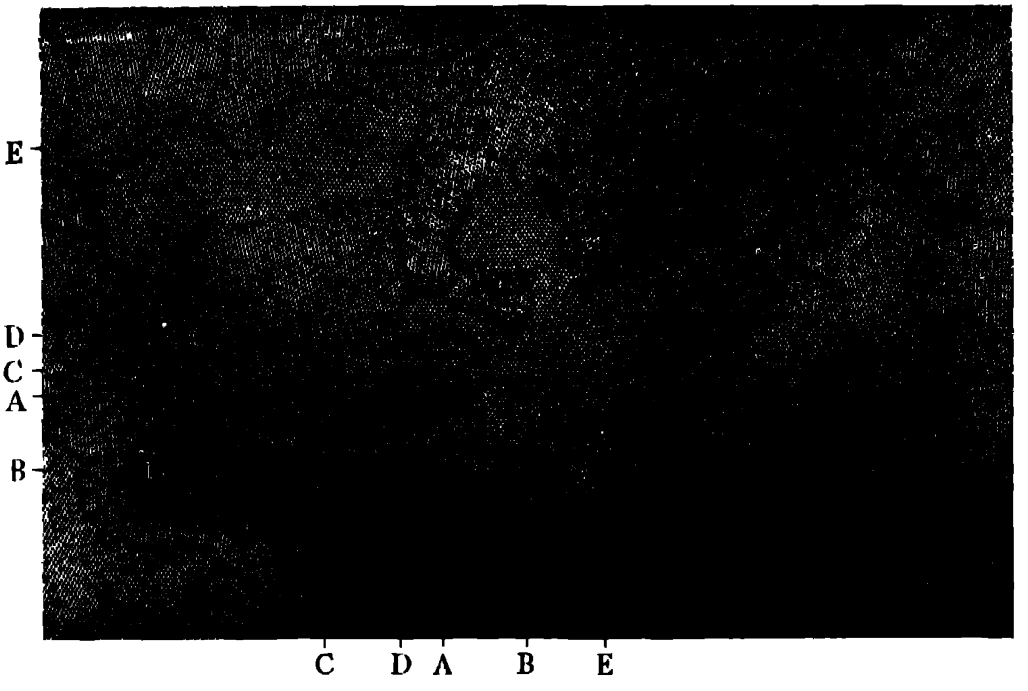


c After 2 min

FIGURE 12

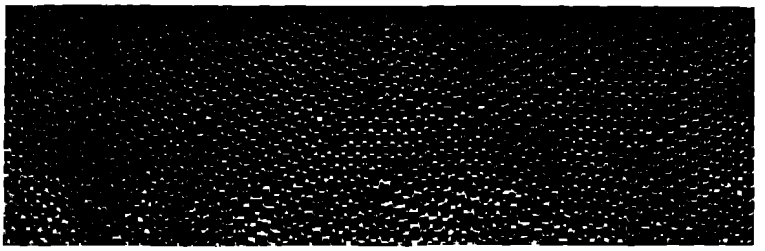


d After 11 mm

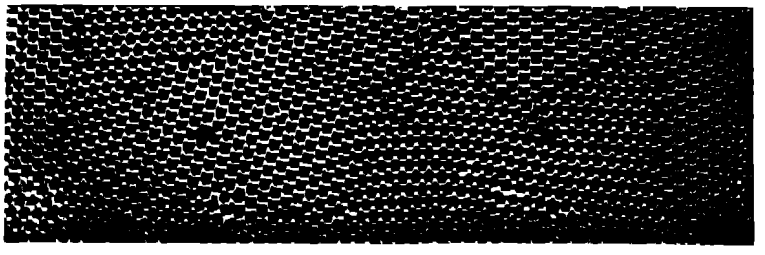


e After 25 mm

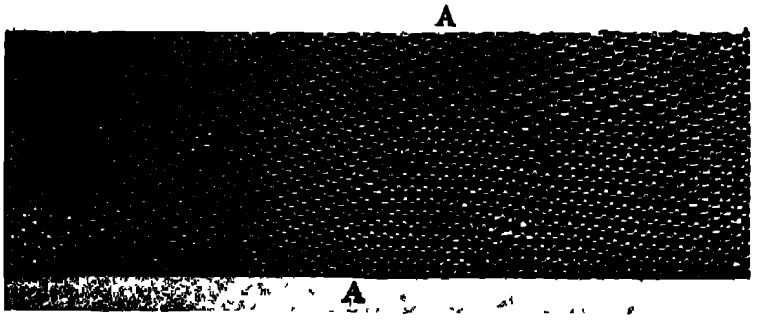
FIGURE 12



a After 1 sec



b After 4 sec



c After 4 min

FIGURE 13 Two stages of recrystallization Diameter 1 64 mm

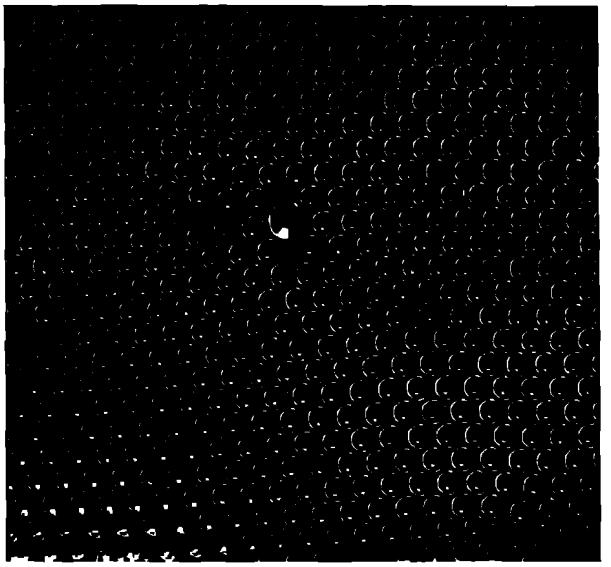
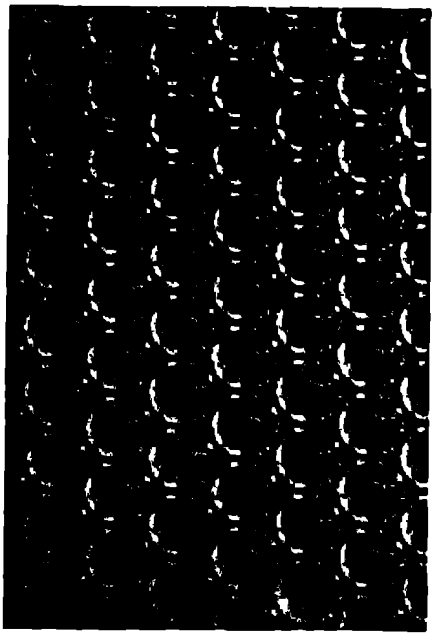


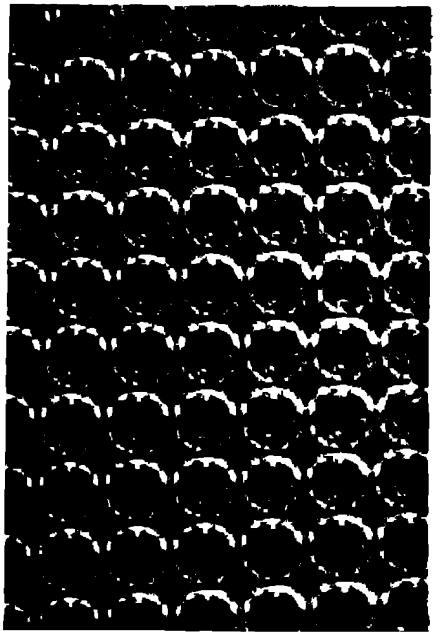
FIGURE 14 Effect of atoms of impurity Diameter of uniform bubbles about 1·3 mm



FIGURE 15 Oblique view of three dimensional raft

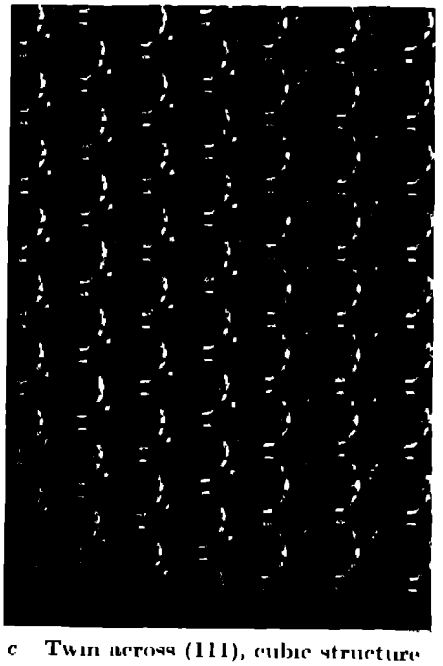


a (111) face

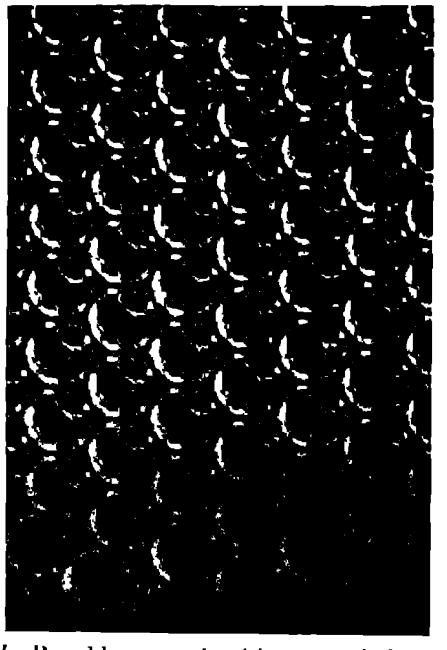


b (100) face

Face-centred cubic structure



c Twin across (111), cubic structure



d Possible example of hexagonal close-packing

Diameter 0·70 mm

FIGURE 17

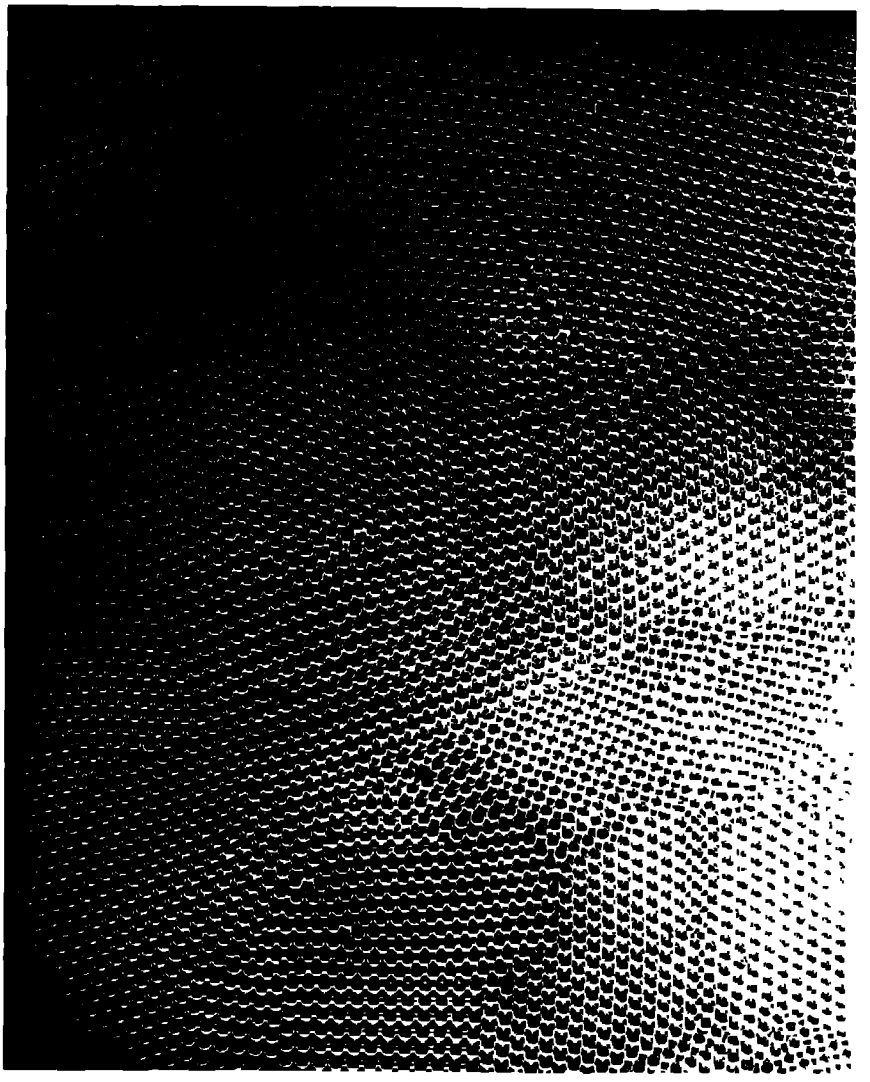


FIGURE 16. A three-dimensional raft viewed normally. Diameter 0·70 mm.

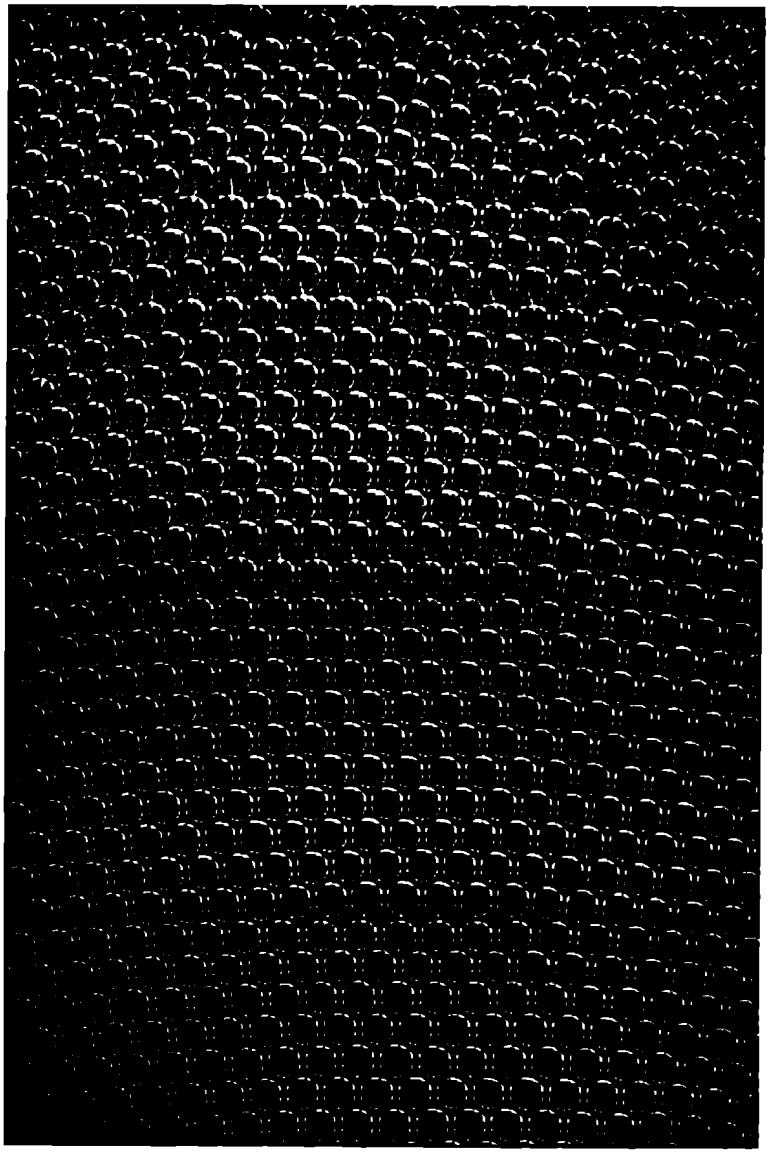


FIGURE 18. Dislocations in three dimensional structure. Diameter 0·70 mm.

much less so for bubbles of 1 mm. diameter, thus confirming the impression given by the model that the small bubbles behave as if they were much more rigid than the large ones.

9. Three-dimensional assemblages

If the bubbles are allowed to accumulate in multiple layers on the surface, they form a mass of three-dimensional 'crystals' with one of the arrangements of closest packing. Figure 15, plate 18, shows an oblique view of such a mass; its resemblance to a polished and etched metal surface is noticeable. In figure 16, plate 20, a similar mass is seen viewed normally Parts of the structure are definitely in cubic closest packing, the outer surface being the (111) face or (100) face Figure 17*a*, plate 19, shows a (111) face. The outlines of the three bubbles on which each upper bubble rests can be clearly seen, and the next layer of these bubbles is faintly visible in a position not beneath the uppermost layer, showing that the packing of the (111) planes has the well-known cubic succession Figure 17*b*, plate 19, shows a (100) face with each bubble resting on four others The cubic axes are of course inclined at 45° to the close-packed rows of the surface layer Figure 17*c*, plate 19, shows a twin in the cubic structure across the face (111) The uppermost faces are (111) and (100), and they make a small angle with each other, though this is not apparent in the figure; it shows up in an oblique view Figure 17*d*, plate 19, appears to show both the cubic and hexagonal succession of closely packed planes, but it is difficult to verify whether the left-hand side follows the true hexagonal close-packed structure because it is not certain that the assemblage had a depth of more than two layers at this point. Many instances of twins, and of intercrystalline boundaries, can be seen in figure 16, plate 20.

Figure 18, plate 21, shows several dislocations in a three-dimensional structure subjected to a bending strain.

10. Demonstration of the model

With the co-operation of Messrs Kodak, a 16 mm cinematograph film has been made of the movements of the dislocations and grain boundaries when single crystal and polycrystalline rafts are sheared, compressed, or extended. Moreover, if the soap solution is placed in a glass vessel with a flat bottom, the model lends itself to projection on a large scale by transmitted light Since a certain depth is required for producing the bubbles, and the solution is rather opaque, it is desirable to make the projection through a glass block resting on the bottom of the vessel and just submerged beneath the surface.

In conclusion, we wish to express our thanks to Mr C E. Harrold, of King's College, Cambridge, who made for us some of the pipettes which were used to produce the bubbles.

References

Bragg, W. L. 1942*a* *Nature*, **149**, 511.
Bragg, W. L. 1942*b* *J. Sci. Instrum.* **19**, 148.
Taylor, G. I. 1934 *Proc. Roy. Soc.* A, **145**, 362.

The accuracy of atomic co-ordinates derived from Fourier series in X-ray structure analysis. III

By A. D Booth, *Physics Department, Birkbeck College*

(*Communicated by G. N. Watson, F R S —Received 2 September* 1946)

Further work on the problems considered in the previous papers of this series has resulted in a more satisfactory treatment of finite summation errors in the three-dimensional diatomic case The results are extended to the two- and one-dimensional series, and the interesting result emerges that finite summation errors are of the same order of magnitude whatever the dimensions of summation

Using the new results a more quantitative examination of the effects of real thermal motion becomes possible It is shown that the relative accuracies of parameters in structures, the higher order reflexions from which are suppressed by thermal motion, follows a simple power law in the corresponding reciprocal spacings

These considerations lead to an examination of the artificial temperature factor method of securing convergence, and it is shown that this produces greater errors due to overlapping than those it is designed to eliminate A method of correcting these distortions is suggested

Finally, the treatment of the effect of experimental errors is extended to two and one dimensions, and it is shown that the three-dimensional summation is least affected by experimental inaccuracy The errors for three-, two- and one-dimensional summation, in a particular case, are calculated to be in the ratio 1:3:10

In previous papers (Booth 1945, 1946, 1947) the author has considered the effects of experimental errors in the Fourier coefficients, and also of the termination of the series, upon the accuracy of atomic co-ordinates derived in the Fourier method of X-ray structure analysis. This paper contains an account of further investigations. An improved treatment of finite summation errors in the diatomic case is given, and a simple relation for comparing the accuracy in structures whose atoms are in different states of thermal motion is derived. The results are applied to an examination of the artificial temperature-factor method of securing convergence

The work is completed by extending all the results to the cases of one- and two-dimensional summation (For §§ 1 to 11 see Booth 1947)

12. Finite summation in the diatomic case

It was shown, in § 6, that the general Fourier summation can be considered to be composed of as many independent parts as there were atoms in the unit cell, and this led, in § 8, to the idea that the finite summation error in the calculated position of a particular atom was the sum of those due to independent incomplete series based on all other atomic positions as origin.

The terminated three-dimensional series for the electron density in the vicinity of an atom at the origin is (6·16)

$$D(x,y,z)=\frac{N}{V}\sum_{-H}^{+H}\sum_{-K}^{+K}\sum_{-L}^{+L}\exp\left[-\frac{\pi^2}{p}\left(\frac{h^2}{a^2}+\frac{k^2}{b^2}+\frac{l^2}{c^2}\right)\right]\cos 2\pi\left(h\frac{x}{a}+k\frac{y}{b}+l\frac{z}{c}\right); \quad (12\cdot1)$$

this may be written approximately

$$D(x,y,z) = \frac{N}{V}\iiint_{\rho} \exp\left[-\frac{\pi^2}{p}\left(\frac{h^2}{a^2}+\frac{k^2}{b^2}+\frac{l^2}{c^2}\right)\right]\cos 2\pi\left(h\frac{x}{a}+k\frac{y}{b}+l\frac{z}{c}\right)\frac{d\xi\,d\zeta\,d\eta}{a^*b^*c^*}, \quad (12{\cdot}2)$$

where (a^*, b^*, c^*) are the reciprocal axes $(\lambda/a, \lambda/b, \lambda/c)$, and the integration extends throughout a sphere in reciprocal space having radius ρ and containing the reciprocal points of all planes included in the summation (12 1) Observing that $h = \xi/a^*$, $k = \zeta/b^*$, $l = \eta/c^*$ the integral (12 2) may be written

$$\begin{aligned} D(x,y,z) &= \frac{N}{\lambda^3}\iiint_{\rho} \exp\left[-\frac{\pi^2}{p\lambda^2}(\xi^2+\zeta^2+\eta^2)\right]\cos\frac{2\pi}{\lambda}(x\xi+y\zeta+z\eta)\,d\xi d\zeta d\eta \\ &= N\iiint_{\sigma} \exp\left[-\frac{\pi^2}{p}(\xi^2+\zeta^2+\eta^2)\right]\cos 2\pi(x\xi+y\zeta+z\eta)\,d\xi d\zeta d\eta, \end{aligned}$$

where the reciprocal region now has radius $\sigma = \rho/\lambda$ Next, let ν be the perpendicular from (ξ, ζ, η) on to the plane

$$x\xi + y\zeta + z\eta = 0;$$

the integral may be transformed to

$$D(s) = N\int_0^{\sigma} 2\pi r \exp\left[-\frac{\pi^2}{p}r^2\right]dr\int_{-r}^{r}\cos(2\pi s\nu)\,d\nu, \quad (12\ 3)$$

where $$s^2 = x^2+y^2+z^2, \quad r^2 = \xi^2+\zeta^2+\eta^2,$$

whence

$$\begin{aligned} D(s) &= \frac{2N}{s}\int_0^{\sigma} r\sin(2\pi sr)\exp\left[-\frac{\pi^2}{p}r^2\right]dr \\ &= -\frac{pN}{\pi^2 s}\exp\left[-\frac{\pi^2}{p}\sigma^2\right]\sin(2\pi s\sigma)+\frac{2pN}{\pi}\int_0^{\sigma}\exp\left[-\frac{\pi^2}{p}r^2\right]\cos(2\pi sr)\,dr. \quad (12{\cdot}4) \end{aligned}$$

Now,
$$\begin{aligned} \frac{\partial D}{\partial s} &= \frac{Np}{\pi}\left(\frac{1}{\pi s^2}\sin(2\pi s\sigma)-\frac{2\sigma}{s}\cos(2\pi s\sigma)\right)\exp\left[-\frac{\pi^2}{p}\sigma^2\right] \\ &\qquad -4pN\int_0^{\sigma} r\exp\left[-\frac{\pi^2}{p}r^2\right]\sin(2\pi sr)\,dr \\ &= \frac{Np}{\pi}\left(\frac{1}{\pi s^2}\sin(2\pi s\sigma)-\frac{2\sigma}{s}\cos(2\pi s\sigma)\right)\exp\left[-\frac{\pi^2}{p}\sigma^2\right]-2psD(s) \quad (12{\cdot}5) \end{aligned}$$

If $\partial D/\partial s$ is to be a maximum $\partial^2 D/\partial s^2 = 0$, i.e.

$$\begin{aligned} &\frac{N}{\pi}\left(-\frac{2}{\pi s^3}\sin(2\pi s\sigma)+\frac{4\sigma}{s^2}\cos(2\pi s\sigma)+\frac{4\pi\sigma^2}{s}\sin(2\pi s\sigma)\right) \\ &\qquad\qquad \times\exp\left[-\frac{\pi^2}{p}\sigma^2\right]-2D-2s\frac{\partial D}{\partial s}=0. \quad (12{\cdot}6) \end{aligned}$$

Whence, substituting for D in (12 5), it is seen that at its maxima

$$\left(\frac{\partial D}{\partial s}\right)_{\max.} = \frac{Np}{\pi^2 s^2}\frac{(3-4\pi^2s^2\sigma^2)\sin(2\pi s\sigma)-6\pi s\sigma\cos(2\pi s\sigma)}{1-2ps^2}\exp\left[-\frac{\pi^2}{p}\sigma^2\right].$$

Now, since these are the values at maxima, it follows that, at *all* points,

$$\frac{\partial D}{\partial s} \leqslant \frac{Np}{\pi^2 s^2} \frac{\sqrt{(16\pi^4 s^4 \sigma^4 + 12\pi^2 s^2 \sigma^2 + 9)}}{1 - 2ps^2} \exp\left[-\frac{\pi^2}{p}\sigma^2\right], \tag{12·7}$$

or, since normally $\pi s\sigma \gg 1$,

$$\frac{\partial D}{\partial s} \leqslant \frac{4Np\sigma^2}{1 - 2ps^2} \exp\left[-\frac{\pi^2}{p}\sigma^2\right]$$

In the case of two-dimensional summation the integral (12 2) becomes

$$D(x, y) = \frac{N}{A} \iint_{\rho} \exp\left[-\frac{\pi^2}{p}\left(\frac{h^2}{a^2} + \frac{k^2}{b^2}\right)\right] \cos 2\pi\left(h\frac{x}{a} + k\frac{y}{b}\right) \frac{d\xi}{a^*} \frac{d\zeta}{b^*}, \tag{12·8}$$

and making similar transformations to the above

$$D(s) = N \int_0^{\sigma} r \exp\left[-\frac{\pi^2}{p} r^2\right] dr \int_{-r}^{r} \frac{\cos(2\pi s\nu)}{\sqrt{(r^2 - \nu^2)}} d\nu,$$

where $$s^2 = x^2 + y^2, \quad r^2 = \xi^2 + \zeta^2.$$

This reduces to

$$D(s) = \frac{pN}{2\pi}\left(1 - \exp\left[-\frac{\pi^2}{p}\sigma^2\right] J_0(2\pi s\sigma)\right) - Nps \int_0^{\sigma} J_1(2\pi sr) \exp\left[-\frac{\pi^2}{p} r^2\right] dr, \dagger$$

and, proceeding exactly as before,

$$\left(\frac{\partial D}{\partial s}\right)_{\max.} = N\, 2\sigma p \frac{(J_1(2\pi s\sigma) - \pi s\sigma J_0(2\pi s\sigma))}{1 - 2ps^2} \exp\left[-\frac{\pi^2}{p}\sigma^2\right]$$

$$= 2Np \frac{\pi s\sigma^2 J_2(2\pi s\sigma)}{1 - 2ps^2} \exp\left[-\frac{\pi^2}{p}\sigma^2\right], \tag{12·9}$$

or approximately $$\frac{\partial D}{\partial s} \leqslant Np\sigma \frac{2\sqrt{(s\sigma)}}{1 - 2ps^2} \exp\left[-\frac{\pi^2}{p}\sigma^2\right].$$

The one-dimensional synthesis can be treated similarly, and the three results are here collected for reference

Three dimensions. $$\frac{\partial D}{\partial s} \leqslant Np \frac{4\sigma^2}{1 - 2ps^2} \exp\left[-\frac{\pi^2}{p}\sigma^2\right]. \tag{12·10}$$

Two dimensions. $$\frac{\partial D}{\partial s} \leqslant Np \frac{2s^{\frac{1}{2}}\sigma^{\frac{1}{2}}}{1 - 2ps^2} \exp\left[-\frac{\pi^2}{p}\sigma^2\right]. \tag{12·11}$$

One dimension. $$\frac{\partial D}{\partial s} \leqslant Np \frac{4s\sigma}{1 - 2ps^2} \exp\left[-\frac{\pi^2}{p}\sigma^2\right], \tag{12·12}$$

† Using the Hankel integral (Watson 1944)

$$\int_0^{\infty} J_{\nu}(at)\, e^{-p^2t^2} t^{\mu-1} = \frac{\Gamma\frac{1}{2}(\mu+\nu)\, a^{\nu}}{\Gamma(\nu+1)\, 2^{\nu+1} p^{\mu+\nu}} {}_1F_1\left(\frac{\mu+\nu}{2}; \nu+1; -\frac{a^2}{4p^2}\right),$$

it is seen that when $\sigma \to \infty$, $D(s) \to N\frac{p}{\pi} e^{-p^2s^2}$, the two-dimensional distribution.

where, in all cases, $\sigma = \rho/\lambda$. It follows from a slight extension of part I, § 4, that the density 'projection' is given by

$$D_n(s) = N\left(\frac{p}{\pi}\right)^{\frac{1}{2}n} e^{-ps^2},$$

where n is the number of dimensions of summation. If a distribution of this form is distant (b) from the origin of the summations considered above, the resulting density is

$$D(s) + D_n(s) = D(s) + N\left(\frac{p}{\pi}\right)^{\frac{1}{2}n} e^{-p(s-b)^2}.$$

The maximum of D_n will be slightly displaced from b, by a small quantity, δ_n say, given by

$$\left(\frac{\partial D(s)}{\partial s}\right)_{s=b} - 2pN\left(\frac{p}{\pi}\right)^{\frac{1}{2}n}\delta_n = 0,$$

neglecting second order terms, whence

$$\delta_n = \frac{1}{2pN}\left(\frac{\pi}{p}\right)^{\frac{1}{2}n}\left(\frac{\partial D}{\partial s}\right)_{s=b}, \tag{12·13}$$

the appropriate expression for the differential coefficient being selected from (12·10) to (12·12).

Assuming the numerical values:

$$p = 4{\cdot}689, \quad S = 1{\cdot}4\,\text{A}, \quad N_1 = N_2, \quad \lambda = 1{\cdot}539\,\text{A},$$

the values of δ_n for different values of ρ have been calculated and are given in table 1.

TABLE 1 TABLE OF δ_n, THE ABSOLUTE ERROR IN TERMINATED 'n'-DIMENSIONAL SUMMATION

ρ ($=2\sin\theta$)	δ_3 (A)	δ_2 (A)	δ_1 (A)
1·2	$0{\cdot}010_7$	$0\,008_7$	$0\,028_6$
1·4	$0\,009_8$	$0{\cdot}006_8$	$0\,021_0$
1 6	$0\,007_0$	$0\,004_9$	$0{\cdot}014_0$
1 8	$0{\cdot}004_8$	$0\,003_2$	$0{\cdot}008_4$
2·0	$0{\cdot}003_0$	$0{\cdot}001_8$	$0\,004_9$
2 2	$0{\cdot}001_7$	$0{\cdot}001_0$	$0\,002_8$

It is seen that the error due to finite termination is of the same order whatever the dimensions of the summation, and also that the more refined treatment given above leads to estimates of error which are in all cases less than those derived by means of the simple considerations of part II, § 8, which are thus justified.

The above method of treatment can of course be extended to the case where the scattering function, here taken as $N\exp\left[-\frac{\pi^2}{p}\left(\frac{h^2}{a^2}+\frac{k^2}{b^2}+\frac{l^2}{c^2}\right)\right]$, is any spherically symmetric function; the results would, however, be the same in the numerical examples examined.

The Fourier integral gives the error as independent of the size of the unit cell. This is, of course, only approximately true, especially in the case of very small cells where the assumption of uniform filling of the reciprocal space implied in (12·2) is not justified. In this case the naive treatment of part II, § 8, would probably give results more in accord with the truth

13 Effect of thermal motion in different structure determinations

It was shown, in (6·12), that the scattering factor, f, of an atom could be represented by

$$f = N \exp\left[-\frac{\pi^2}{p}\left(\frac{\rho}{\lambda}\right)^2\right], \tag{13·1}$$

and since the limit of observable $|F|$ values, in similar structures containing the same number of atoms, is directly proportional to f, it follows that if in the two structures the observable reflexions extend to reciprocal spacings ρ_1 and ρ_2 respectively, then

$$f_{\lim} = N_1 \exp\left[-\frac{\pi^2}{p_1}\left(\frac{\rho_1}{\lambda}\right)^2\right] = N_2 \exp\left[-\frac{\pi^2}{p_2}\left(\frac{\rho_2}{\lambda}\right)^2\right],$$

or, if the structures contain the same type of atom,

$$\frac{\rho_1^2}{p_1} = \frac{\rho_2^2}{p_2}. \tag{13·2}$$

Now the errors, in each case, are given by (12·10) to (12·13), whence, making various obvious approximations and using (13·2), it follows that

$$\frac{{}_1\delta_n}{{}_2\delta_n} = \left(\frac{\rho_2}{\rho_1}\right)^{k(3+n)}, \tag{13·3}$$

where n is the number of dimensions of summation. If the structures contain widely differing numbers of atoms the above simple treatment is not valid and it would be necessary to work with the values of $|F|$ instead of f This, unfortunately, renders the direct comparison given above impossible, and up to the present no reasonable method of simplification has been found

14 Improvement of co-ordinates

The computation of correcting syntheses as described in part II, § 10, is necessarily a laborious operation, and the author was led to examine the 'artificial temperature factor' (suggested by Bragg & West (1930) and van Reijen (1942) as an aid to removing spurious detail in the interatomic regions of a synthesis) as a means of reducing finite summation errors The introduction of this factor is made by multiplying each term of a synthesis by the appropriate value of $\exp -\alpha(\rho/\lambda)^2$, where α is a constant The result is to make all high-order terms very small and so improve the convergence of the series The effect must not be confused with that considered in § 12, since, in the present case, the high-order terms *are* observed and *are* included on a very small scale in the synthesis.

A synthesis, modified by the introduction of a factor of the type mentioned, will be characterized by the same atomic positions but different atomic shapes, the scattering factor of an atom in the new distribution being

$$f^* = f\exp\left[-\alpha\left(\frac{\rho}{\lambda}\right)^2\right],$$

or, substituting from (13·1) for f,

$$f^* = N\exp\left[-\pi^2\left(\frac{1}{p}+\frac{\alpha}{\pi^2}\right)\left(\frac{\rho}{\lambda}\right)^2\right]. \tag{14·1}$$

The density distribution corresponding to this scattering factor can be obtained, either by Fourier inversion, or, more simply, by comparison with (4·2) and (6·12), and is

$$N\left(\frac{p\pi}{\pi^2+p\alpha}\right)^{\frac{3}{2}}\exp\left[-\frac{p\pi^2}{\pi^2+p\alpha}r^2\right]. \tag{14 2}$$

Now it is evident from (12 10) to (12·13), that by making α large enough the finite summation error can be made as small as desired, but at the same time the spreading of the density distribution, implied in (14 2), will cause mutual distortion of the atomic peaks in the manner treated in part II, § 7. Using (14·2) with (7 4), the deviation produced in each of a pair of adjacent atomic peaks is

$$\Delta = \frac{N_2}{N_1}s\exp\left[-\frac{p\pi^2}{\pi^2+p\alpha}s^2\right] \tag{14 3}$$

The interesting function is the sum of the two errors $\Delta+\delta$ Modifying (12·10) in the light of (14 1), δ becomes (in the three-dimensional case)

$$\left.\begin{aligned}\delta_3 &= \frac{N_2}{N_1}\left(\frac{\pi}{p^*}\right)^{\frac{3}{2}}\frac{2\sigma^2}{1-2p^*s^2}\exp\left[-\frac{\pi^2}{p}\sigma^2\right]\exp\left[-\alpha\sigma^2\right],\\ \text{where}\quad p^* &= \frac{p\pi^2}{\pi^2+p\alpha}\end{aligned}\right\} \tag{14 4}$$

The values of these functions are given in table 2 for various values of α, it is assumed that $\rho = 1\ 8$, as is usual with Cu K_α radiation, $S = 1\ 4$ A, and $N_1 = N_2$

TABLE 2 SHOWING VARIATION OF TOTAL CO-ORDINATE ERROR WITH APPLIED TEMPERATURE FACTOR

α = temperature factor	δ_3 = error due to finite summation (A)	Δ = error due to overlapping (A)	$\Delta+\delta_3$ (A)
0·0	0·0048	0·0001	0·0049
0 5	0 0045	0·0008	0·0053
1 0	0 0033	0 0028	0·0061
1 5	0 0025	0 0066	0 0091
2·0	0·0018	0 0128	0 0146
3 0	0 0008	0·0315	0·0323

It follows that as a single means of increasing accuracy the artificial temperature-factor method is untenable Its possible use in this field would be, however, to eliminate the finite summation error almost entirely by taking $\alpha > 3\ 0$ say, and then

to correct the resulting atomic positions for mutual interference by means of (14·3). Possible advantages of this method lie in the facts that atoms, situated symmetrically with respect to their neighbours, require no correction, and also, that the effect of the experimental errors, which are most pronounced in small high-order terms, is diminished. In addition, the effect of closest neighbours need alone be considered, since the more remote atoms will contribute insignificantly to the distribution.

15. Experimental errors in one- and two-dimensional syntheses

In the case of a one-dimensional or line synthesis the distribution function is

$$N\sqrt{\frac{p}{\pi}}e^{-px^2}, \tag{15·1}$$

having the Fourier expansion

$$\frac{1}{a}\sum_{-H}^{H}|F|\cos\left(2\pi h\frac{x}{a}-\alpha\right). \tag{15·2}$$

Proceeding in the same manner as in part I, § 5, the error in co-ordinates, due to experimental errors in the $|F|$, is found to be given by

$$\epsilon_x<\frac{1}{Na^2}\left(\frac{\pi}{p}\right)^{\frac{1}{2}}\left(\sum_{-H}^{H}h^2\right)^{\frac{1}{2}}\Delta e, \tag{15·3}$$

or, if H is the highest observed order of reflexion,

$$\epsilon_x<\frac{1}{Na^2}\left(\frac{\pi}{p}\right)^{\frac{1}{2}}\left[\frac{H}{3}(H+1)(2H+1)\right]^{\frac{1}{2}}\Delta e.$$

Assuming H to be fairly large and observing that $H=2a/\lambda$, this becomes

$$\epsilon_x<\frac{4}{N\sqrt{(3a)}}\left(\frac{\pi}{\lambda p}\right)^{\frac{1}{2}}\Delta e \tag{15·4}$$

And taking, as in part I,

$$N=6,\quad a=6{\cdot}150\,\text{A},\quad \lambda=1{\cdot}539\,\text{A},\quad p=4{\cdot}689,\quad \Delta e=0{\cdot}6,$$

(15·4) gives $$\epsilon_x<0{\cdot}027\,\text{A},$$

or ten times the error produced in three-dimensional synthesis.

In the two-dimensional case the distribution function is

$$N\left(\frac{p}{\pi}\right)e^{-p(x^2+y^2)}, \tag{15·5}$$

and the error is given by $$\epsilon_x<\frac{1}{Na^2b}\left(\frac{\pi}{p}\right)^2\left(\sum_{-H}^{+H}\sum_{-K}^{+K}h^2\right)^{\frac{1}{2}}\Delta e, \tag{15·6}$$

the series can be summed approximately as an integral, and the error turns out to be

$$\epsilon_x<\frac{2}{N}\sqrt{\frac{\pi}{ab}}\left(\frac{\pi}{p\lambda}\right)^2\Delta e. \tag{15·7}$$

For the special case given above, with the addition of $b=12{\cdot}312\,\text{A}$, the numerical value is

$$\epsilon_x<0{\cdot}008\,\text{A},$$

again larger than the error for three-dimensional summation.

These results form the *raison d'être* of three-dimensional syntheses, whose merit is seen to be the reduction of experimental errors. The three forms are collected together for reference below:

One-dimensional synthesis. $$\epsilon_x < \frac{2{\cdot}31}{N\sqrt{a}}\left(\frac{\pi}{\lambda p}\right)^{\frac{1}{2}} \Delta e$$

Two-dimensional synthesis. $$\epsilon_x = \epsilon_y < \frac{3{\cdot}54}{N\sqrt{A}}\left(\frac{\pi}{\lambda p}\right)^{2} \Delta e$$

Three-dimensional synthesis:

$$\epsilon_x = \epsilon_y = \epsilon_z < \frac{5{\cdot}18}{N\sqrt{V}}\left(\frac{\pi}{\lambda p}\right)^{\frac{1}{2}} \Delta e$$

It is interesting to note that if the value $p = 4\ 689$ be taken these results take the form, for Cu K_α radiation,

$$\epsilon < \frac{0\ 66 \Delta e}{N\sqrt{v}}, \tag{15·8}$$

where v represents a, A or V according to the dimensions of the synthesis.

16. Conclusion

A more refined treatment of the question of deviations of maxima due to finite summation in Fourier synthesis leads to the result that errors due to this cause are of the same order whatever the dimensions of the summation, assuming, of course, that the resolution is the same in each case The superiority of the three-dimensional methods is revealed in an examination of the co-ordinate deviations produced by experimental errors in the $|F|$ values, where examination of a numerical example shows the one-dimensional synthesis to introduce errors ten times as great as those in a corresponding three-dimensional synthesis.

A critical discussion of the 'artificial temperature factor' method of securing rapid convergence shows it to result in large errors due to mutual interference of electron density distributions and a method of correcting for this displacement is suggested.

Quantitative examination of the problem of real thermal motion shows the errors to increase with increasing thermal vibration, in accordance with a simple power law

The author expresses his gratitude to Professor G N Watson, F.R S., for his continued interest and encouragement, and to Dr K Lonsdale, F.R.S , for much appreciated criticism.

References

Booth 1945 *Phil Mag.* **36**, 609.
Booth 1946 *Trans. Faraday Soc.* **42**, 444.
Booth 1947 *Proc Roy. Soc.* A, **188**, 77.
Bragg & West 1930 *Phil Mag.* **10**, 823
van Reijen 1942 *Physica, Eindhoven*, **9**, 461.
Watson, G N. 1944 *Theory of Bessel functions* Camb. Univ. Press.

The accuracy of atomic co-ordinates derived from Fourier series in X-ray structure analysis

IV. The two-dimensional projection of oxalic acid

By A. D Booth, *Physics Department, Birkbeck College*

(*Communicated by J M Robertson, F R.S —Received* 30 *October* 1946)

The methods, developed in previous papers of this series, have been applied to an examination of the errors in the atomic co-ordinates derived from the (h, 0, l) projection of oxalic acid dihydrate.

It is shown that the experimental errors are of the order ± 0.01 A, and that the finite summation errors are slightly larger—in agreement with the theoretical predictions of the former papers

Recent work, using more extensive experimental data, is discussed, and it is concluded that, owing to the introduction of an artificial temperature factor, the results are unlikely to be of greater accuracy than the originals.

The previous papers of this series (Booth 1945, 1946, 1947*a*, *b*) have been directed towards an examination of the general theory underlying the Fourier series used in X-ray crystal structure analysis and, except where necessary to evaluate numerical constants, no specific reference has been made to actual structure determinations In this, and succeeding papers, the results of the theoretical treatment will be applied to the examination of particular structure analyses. In view of the not inconsiderable labour required for the computations, the structures chosen for detailed examination are not only interesting from the point of view of the general theory of accuracy, but also from their importance in the theory of chemical bonding

Of such structures, one of the first to receive detailed examination by X-ray methods was that of oxalic acid-dihydrate An approximate determination was made by Zachariasen (1934), and this was followed by a complete two-dimensional Fourier refinement by Robertson & Woodward (1936) Interest immediately centred on the structure on account of the short bridges *ab* (figure 1) These remain, to the present time, the shortest known hydrogen bridge between oxygen atoms (Robertson 1947)

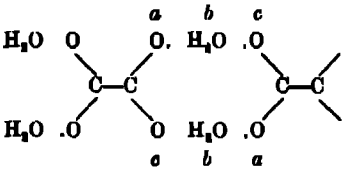


Figure 1

A recent re-examination of the structure by Brill, Hermann & Peters (1939, 1942), using all two-dimensional data obtainable with molybdenum radiation, claimed extreme accuracy and disagreed in important details with the Robertson-Woodward

results. It is unfortunate that the two-dimensional technique, when applied to this compound, breaks down owing to the non-resolution of the (*a*) and (*c*) projections and, despite elaborate corrections by Brill *et al*, it cannot be said that the structure is unequivocally established as a result of their work. Although none of the methods developed in the present papers is applicable to the case of non-resolved projections it seemed desirable to apply the techniques to the projection on (*b*), as excellent resolution is obtained and interest attaches to the considerable disagreement between the co-ordinates of the carbon atom, as deduced by Robertson & Woodward, and given by Brill *et al.*

Experimental errors

Just as a comparison between the Robertson and the Jeffrey observed structure factors for dibenzyl showed a normal distribution of errors (Booth 1947*a*), and enabled the probable error in atomic co-ordinates to be calculated from the equation

$$\epsilon \leqslant \pm \frac{0{\cdot}66\Delta e}{N\sqrt{v}}, \tag{1}$$

where ϵ = probable error in co-ordinate, in A, Δe = probable error in experimental value of $F(h,k,l)$, N = atomic number of particular atom, v = 'volume' of repeat unit, so, in the case of oxalic acid, a comparison between the data obtained by Robertson & Woodward, and by Brill *et al*, enables the experimental error in the $F(h,k,l)$ values to be estimated. In figure 2 is shown the distribution of errors obtained from this comparison By the usual methods the most probable error is found from this curve to be

$$\Delta e = \pm 0{\cdot}82. \tag{2}$$

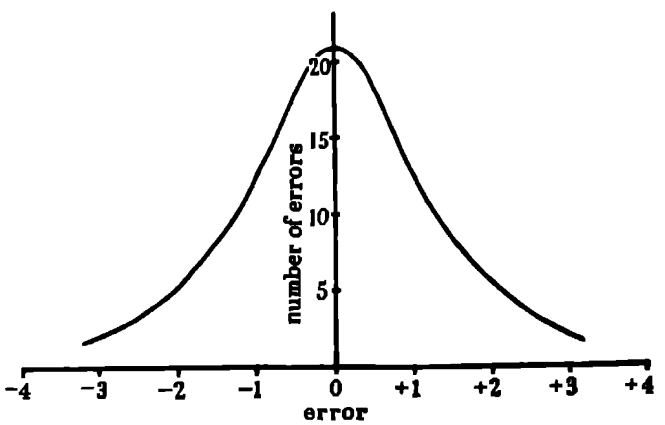


Figure 2

This is in satisfactory accord with the value $\pm 0{\cdot}6$ obtained in the dibenzyl comparison. The generalized volume, 'v', is simply the area of the projection repeat,

in the case of oxalic acid 70·5 A². Thus, for the carbon atoms, where $N = 6$, substitution in equation (1) gives

$$\epsilon \leqslant \pm 0{\cdot}010_7 \text{ A},$$

and for the oxygen atoms, where $N = 8$,

$$\epsilon \leqslant \pm 0{\cdot}008 \text{ A}.$$

These values are of precisely that order of magnitude which would be expected after the previous work on dibenzyl.

Errors due to finite summation

In order to apply the correction technique suggested in part II of this paper (Booth 1947*a*), it was necessary to have a method of computation of greater accuracy than those in current use. This led to the development of the differential synthesis, which makes possible the use of three- or four-figure mathematical tables, and reduces errors of computation to *c* ±0 001 A (Booth 1946) In the case of oxalic acid it was not deemed advisable to assume that the electron density projections, on the (*b*) plane, were possessed of circular symmetry, so that all coefficients in the difference equations (vide infra) were calculated

The electron density projection, for oxalic acid, is given by

$$\rho(x,z) = \frac{1}{A}\sum_{-\infty}^{+\infty}_{h}\sum_{-\infty}^{+\infty}_{l} F(h,0,l)\cos 2\pi\left(h\frac{x}{a}+l\frac{z}{c}\right), \tag{3}$$

where $\rho(x,z)$ is the electron density per square A, A is the area of the repeat unit, and $F(h,0,l)$ is the structure factor If the point (x,z) coincides with a maximum of electron density (x_0,z_0) it follows that

$$\left(\frac{\partial\rho}{\partial x}\right)_0 = \left(\frac{\partial\rho}{\partial z}\right)_0 = 0. \tag{4}$$

These relations, in conjunction with equation (3), give

$$\sum_{0}^{+\infty}\sum_{0}^{+\infty} -2\pi\frac{h}{a}\left\{F(h,0,l)\sin 2\pi\left(h\frac{x_0}{a}+l\frac{z_0}{c}\right)+F(h,0,\bar{l})\sin 2\pi\left(h\frac{x_0}{a}-l\frac{z_0}{c}\right)\right\} = 0, \tag{5}$$

$$\sum_{0}^{+\infty}\sum_{0}^{+\infty} -2\pi\frac{l}{c}\left\{F(h,0,l)\sin 2\pi\left(h\frac{x_0}{a}+l\frac{z_0}{c}\right)-F(h,0,\bar{l})\sin 2\pi\left(h\frac{x_0}{a}-l\frac{z_0}{c}\right)\right\} = 0. \tag{6}$$

Now, if the true co-ordinates differ from (x_0,z_0) by small quantities δx and δz, substitution in (5) and (6), and rejection of second and higher powers of small quantities, leads to the equations

$$A_{hh}\frac{\delta x}{a}+A_{hl}\frac{\delta z}{c}+A_h = 0, \tag{7}$$

$$A_{lh}\frac{\delta x}{a}+A_{ll}\frac{\delta z}{c}+A_l = 0, \tag{8}$$

where $A_{hl} = A_{lh}$ and

$$A_{hh} = 2\pi \sum_{0}^{+\infty} \sum_{0}^{+\infty} h^2 \left\{ F(h,0,l) \cos 2\pi \left(h\frac{x_0}{a} + l\frac{z_0}{c} \right) + F(h,0,\bar{l}) \cos 2\pi \left(h\frac{x_0}{a} - l\frac{z_0}{c} \right) \right\} = 0, \quad (9)$$

$$A_{hl} = 2\pi \sum_{0}^{+\infty} \sum_{0}^{+\infty} hl \left\{ F(h,0,l) \cos 2\pi \left(h\frac{x_0}{a} + l\frac{z_0}{c} \right) - F(h,0,\bar{l}) \cos 2\pi \left(h\frac{x_0}{a} - l\frac{z_0}{c} \right) \right\} = 0, \quad (10)$$

$$A_h = \sum_{0}^{+\infty} \sum_{0}^{+\infty} h \left\{ F(h,0,l) \sin 2\pi \left(h\frac{x_0}{a} + l\frac{z_0}{c} \right) + F(h,0,\bar{l}) \sin 2\pi \left(h\frac{x_0}{a} - l\frac{z_0}{c} \right) \right\} = 0, \quad (11)$$

$$A_l = \sum_{0}^{+\infty} \sum_{0}^{+\infty} l \left\{ F(h,0,l) \sin 2\pi \left(h\frac{x_0}{a} + l\frac{z_0}{c} \right) - F(h,0,\bar{l}) \sin 2\pi \left(h\frac{x_0}{a} - l\frac{z_0}{c} \right) \right\} = 0. \quad (12)$$

In the present case, the co-ordinates of Robertson & Woodward were taken as the first approximation, and, using the structure factors *calculated* from these co-ordinates, a set of differential syntheses was calculated for each atom. The method of computation which was found most convenient was to draw up tables of

$$\sin 2\pi \left(h\frac{x_0}{a} \pm l\frac{z_0}{c} \right) \quad \text{and} \quad \cos 2\pi \left(h\frac{x_0}{a} \pm l\frac{z_0}{c} \right),$$

using a Marchant multiplying machine, and then to multiply by the values of $F(h,0,l)$ The coefficients A_{hh}, etc., were then easily obtained Table 1 gives the values of the various coefficients From these values the errors in co-ordinates, due to finite summation, can be calculated via equations (7) and (8) The *corrections* to the co-ordinates are obtained by reversing the signs of these errors, they are given in table 2

TABLE 1

atom	A_{hh}	A_{hl}	A_{ll}	A_h	A_l
C	+2117 9	− 804·4	+ 8,565 2	−32 8	+45·8
O_1	+2884 5	−1014·1	+11,861·7	+13 2	+96·0
O_2	+3133·3	−1586·7	+10,350·1	+40·2	+ 2 6
H_2O	+2715·2	−1603·4	+10,912 0	−60·8	−23·4

TABLE 2

atom	$\delta x/a$	$\delta z/c$
C	−0·0022	+0 0007
O_1	+0·0012	+0·0014
O_2	+0·0022	+0·0003
H_2O	−0·0041	−0·0009

In order to guard against computing errors a set of second-order refinements using the corrected co-ordinates was calculated. It was not necessary to recalculate A_{hh}, A_{hl} and A_{ll}, since these are relatively insensitive to changes in the co-ordinates. The important factors are A_h and A_l, and the values obtained by using the corrected co-ordinates are given in table 3.

TABLE 3

atom	$\delta x/a$	$\delta z/c$
C	+0·0000	−0 0001
O_1	+0·0002	−0·0000
O_2	+0 0000	−0 0000
H_2O	+0·0002	+0 0004

These figures indicate the correctness of the original values and establish the validity of the differential method It is seen that the second-order corrections are roughly one-tenth as large as those of the first order. Table 4 gives the corrected atomic co-ordinates together with those found by Brill *et al*

TABLE 4

	R and W corr.		B, H and P obs.	
atom	x/a	z/c	x/a	z/c
C	−0 0389	+0 0514	−0 0465	+0 0511
O_1	+0 0872	+0·1492	+0 0828	+0 1478
O_2	−0·2258	+0·0353	−0 2192	+0 0362
H_2O	−0 4559	+0 1784	−0 4549	+0 1780

It is seen that there are still considerable differences between the two sets, although the method of correction gives results which are usually of the same sign as those observed by Brill *et al*

The greatest discrepancy is in the x co-ordinate of C, it was thus of interest to examine the method by which the result of Brill *et al* was obtained Despite the fact that the series converged to the limit of the observable $F(h, k, l)$ values these authors applied the artificial temperature-factor technique to reduce possible finite summation errors This, as was demonstrated in part III of this series, introduces large errors due to overlapping of the density distributions, realizing this, corrections were made by Brill *et al.* for this error To investigate the effectiveness of these corrections the same process was applied to the present calculations.

A temperature factor $e^{-5\alpha(\sin\theta)^2}$ (13)

was applied to the values of $F(h, 0, l)$, and a differential synthesis computed for the atom C The coefficients obtained were

$$A_{hh} = +\ 212{\cdot}4, \qquad A_h = +28{\cdot}3,$$
$$A_{hl} = -\ 219{\cdot}5, \qquad A_l = -87{\cdot}7.$$
$$A_{ll} = +1094\ 3,$$

These lead to distortions

$$\delta x/a = -0{\cdot}0103, \qquad \delta z/c = +0{\cdot}0107,$$

which are of approximately ten times the magnitude of the original finite summation errors given in table 2 Applying corrections for the overlapping distortion reduces these values to

$$\delta x/a = -0{\cdot}0003, \qquad \delta z/c = +0{\cdot}0068,$$

but even so the method does not appear to be as satisfactory as the technique used in this paper It was not considered worth applying the procedure to the other atoms in the structure.

A possible criticism of the method of *ad hoc* synthesis using calculated values of $F(h, 0, l)$, is that the differences between observed and calculated F values produce exactly the same variation in co-ordinates as those obtained in the correction process In order to justify the procedure it is necessary to show that the difference between F_{calc} and F_{obs} produces approximately the same background undulation whatever the maximum value of $\sin\theta$ at the limit of summation. To test this, a series of computations was made using data extending to various Bragg angles. It is clear from equations (5), (6), (11) and (12), that

$$A_h \propto \frac{\partial \rho}{\partial x}, \quad A_l \propto \frac{\partial \rho}{\partial z},$$

so that a calculation of these quantities gives a measure of the background fluctuation

The atom C was again chosen for the test, and the resulting coefficients are given in table 5. It is seen that, whilst there is considerable fluctuation in the finite summation differential, the slope of the background given by the $(F_{calc} - F_{obs})$ set, is approximately constant.

TABLE 5

	using (F_{calc})		using $(F_{calc} - F_{obs})$	
$2 \sin\theta$	A_h	A_l	A_h	A_l
1 8	−32·8	+45 8	−45 8	−86 8
1 6	− 0·8	+90 6	−43 8	−63 6
1 4	−30 2	+15 8	−55 0	−58 6
1 2	+19 8	+60·8	−57 6	−40 8
1 0	−17·7	−95·7	−44 2	−35 4

Whilst it cannot be claimed that the foregoing treatment constitutes definite proof of the validity of the correction technique, it indicates that the errors derived from it are at least of the right order of magnitude.

The results obtained by Brill *et al.* for oxalic acid have been questioned on grounds of chemical plausibility by Robertson (1947), who shows that the bond lengths deduced on the basis of the present calculations, are more satisfactory than those of the former authors At the same time, the unsatisfactory resolution of the other projections makes any final decision impossible

The usual criterion (Booth 1945) of the degree of excellence of a set of atomic parameters is the value of the expression

$$\frac{\Sigma \,||\, F_{obs} \,| - |\, F_{calc} \,||}{\Sigma \,|\, F_{obs} \,|}. \qquad (14)$$

In the case of oxalic acid the values of this expression for the various determinations which have been made are

Zachariasen	0·156
Robertson & Woodward	0·124
Brill *et al.*	0·110
Robertson & Woodward with corrections	0·113
Robertson re-examination	0·102

It is seen that the corrections derived above improve the value, and are nearly as effective in this direction as the Brill *et al.* co-ordinates The last figure is derived from a structure proposed by Robertson (1947) which is particularly satisfactory on chemical grounds

The value of (14) is not an unequivocal guide to the excellence of a set of parameters as it depends, not only upon the correctness of the co-ordinates and the experimental errors, but also on the departure of the electron distribution from spherical symmetry which, in a covalently linked system, may be considerable Thus the only manner in which the correctness of the various alternatives presented above can be established, is by performing a full three-dimensional analysis and using all information obtainable with molybdenum or shorter wave-length radiation, but avoiding the use of the artificial thermal factor

References

Booth 1945 *Phil. Mag* **36**, 609.
Booth 1946 *Trans. Faraday Soc* **42**, 444.
Booth 1947*a* *Proc. Roy. Soc* A, **188**, 77.
Booth 1947*b* *Proc. Roy Soc.* A, **190**, 482
Brill, Hermann & Peters 1939 *Naturwissenschaften*, **27**, 677.
Brill, Hermann & Peters 1942 *Ann. Phys., Lpz.*, **42**, 357
Robertson 1947 *J Chem Soc* (in the Press)
Robertson & Woodward 1936 *J. Chem Soc* p. 1817.
Zachariasen 1934 *Z. Kristallogr.* **9**, 442

Local and extensive penetrating cosmic ray showers

By D. Broadbent and L Jánossy*

The Physical Laboratories, The University, Manchester

(*Communicated by P M. S Blackett, F R S.—Received 29 November* 1946)

Penetrating cosmic ray showers are observed with an anticoincidence method which indicates whether or not each shower is associated with an extensive air shower Those not associated with extensive showers (local penetrating showers) are found to show a transition effect which is mass-proportional, while the others (extensive penetrating showers) show a transition effect which is strongly Z-dependent, and is very like the Rossi transition effect of cascade showers It is concluded that these are two completely independent types of penetrating showers, the local showers being groups of mesons produced in absorbers near the apparatus by single fast primary nucleons, and the extensive penetrating showers being large cascade showers containing penetrating particles This separation is shown to accord well with previous observations

1 Introduction

In a number of investigations Jánossy and his co-workers have investigated the properties of penetrating showers Some of the penetrating showers were found to be associated with extensive air showers, and the present paper deals mainly with this connexion. A penetrating shower set P was used together with an anti-coincidence counter-tray E placed near P which indicated whether or not any penetrating shower observed was associated with an extensive shower It is shown that the penetrating showers are of two types (a) penetrating showers not associated with extensive showers (local penetrating showers), and (b) penetrating showers associated with extensive air showers (extensive penetrating showers). These two types show quite different properties and are probably completely independent of each other. The fact that the two types of showers give rise to roughly the same rates of coincidences for the arrangements employed previously must be taken as accidental; this circumstance gave rise, however, to much confusion in the past.

2. The experimental procedure

The experimental arrangement consisted of a penetrating shower set P very similar to that used on previous occasions and described fully by Jánossy (1942) and by Jánossy & Rochester (1944) The set P (figure 1) contains three trays of counters t, m, b. Each of the trays contains eight counters and each tray covers an area of 1000 cm ^{2}. The trays t, m and b are separated by 15 cm. of lead absorbers, and the trays m and b shielded at the sides by about 50 cm of lead The top tray t is unshielded and is split into three groups t_1, t_2, t_3, two of these containing three counters

* Now at the Dublin Institute for Advanced Studies.

each and the third containing only two counters. The trays m and b are each split into two groups of four counters, which may be denoted m_1, m_2, b_1, b_2. Sevenfold coincidences (t_1, t_2, t_3, m_1, m_2, b_1, b_2) are recorded, and such coincidences will be denoted shortly as coincidences (P). It has been shown previously that coincidences (P) are mainly caused by penetrating showers different from ordinary cascades or from knock-on showers.

In the present experiments a large tray E containing twelve counters in parallel was placed near to the tray t. The tray E covered 2500 cm ^{2}, and the centres of the t and E trays were about 50 cm apart Both eightfold coincidences (P, E) and anti-coincidences (P–E) were recorded Anticoincidences (P–E) were the coincidences of P not accompanied by the discharge of any of the counters E.

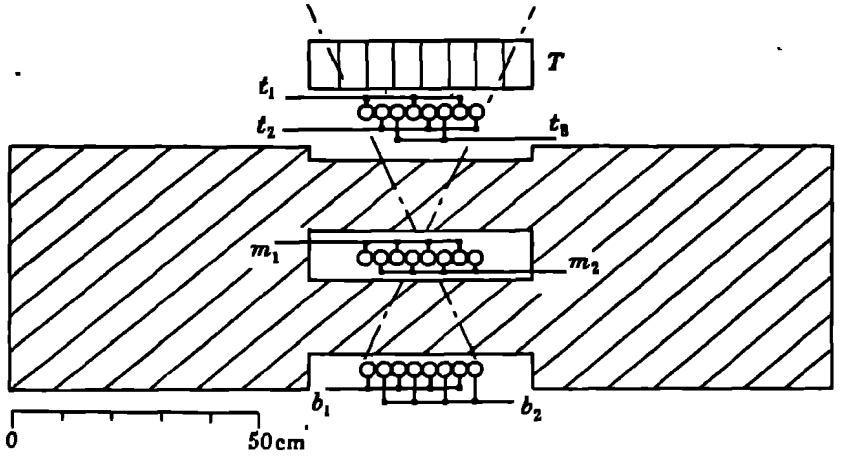


FIGURE 1 The penetrating shower set P

An upper limit to the rate of accidental coincidences (P, E) was obtained by measuring twofold coincidences between E and the set t_1. This twofold rate was about 0·5 % of the single rate of t_1, and because of the high single rates of E and t_1—about 10,000 and 1000 counts/min respectively—was due almost entirely to accidental coincidences The rate of accidental coincidences (P, E) amounted thus to not more than 0·5 % of the rate of coincidences (P). In view of the very low rate of the coincidences (P) this accidental rate could be neglected throughout the experiments, it corresponded to a resolving time for the coincidence circuit of the order of 10^{-5} sec

The twelve counters E were used with a Neher-Harper circuit, and to reduce inefficiency the counters were split into four groups, each group having a separate quenching valve The rate of spurious anticoincidences due to inefficiency was estimated to be about 1 % of the rate of genuine anticoincidences This did not affect appreciably the rates of anticoincidences (P–E), and no correction for it was found to be necessary.

3. The experimental results

3·1. *Anticoincidence records*

Anticoincidences (P–E) were recorded with various absorbers T close above the counters t The absorbers covered an area of 45×45 cm.^{2}, and readings were taken with lead and with paraffin absorbers. Two thicknesses of paraffin, 14 and 28 cm. were used, corresponding on a mass basis to the 1 and 2 cm thicknesses of lead. To make a strict comparison between the lead and paraffin these lead absorbers were made up of sheets about 3 mm. thick spaced to extend over the same depth as was occupied by the corresponding paraffin blocks.

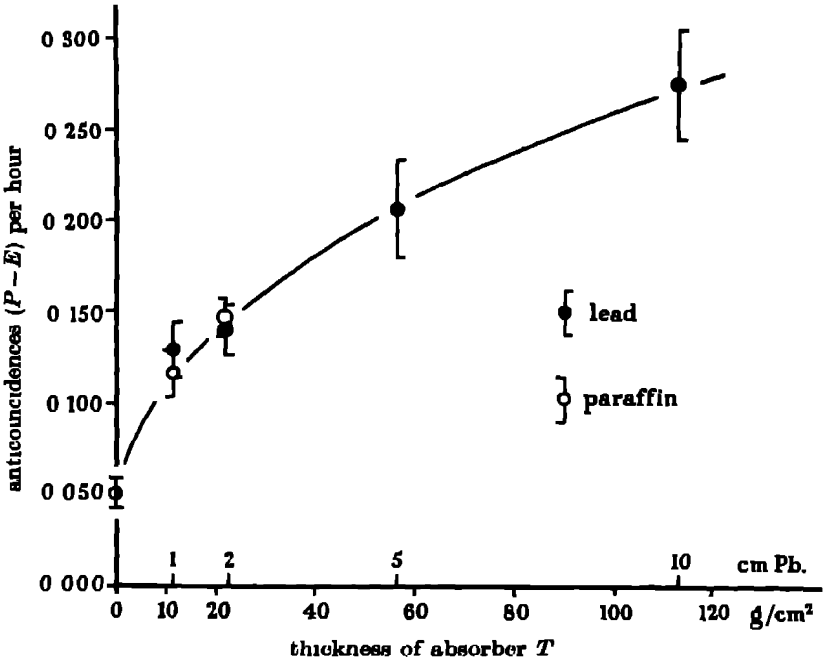


Figure 2. Transition effect of local penetrating showers.

For reasons of geometry it did not appear useful to increase the paraffin absorbers beyond 28 cm , for greater thicknesses lead absorbers only were used, and these thicker absorbers were made up of solid blocks.

The rates of anticoincidences (P–E) are collected in table 1 and represented graphically in figure 2, the errors quoted being the standard deviations The total time of recording involved in the observations of table 1 was about 4000 hr.

It is seen from table 1 that the rate of anticoincidences for $T = 0$ is very small, and that the rate increases quite considerably with increasing T The rate increases about threefold for both 22·6 g./cm.^{2} lead and paraffin. Saturation is only reached for large thicknesses The rate under 10 cm. of lead is about five times the zero rate

TABLE 1. RATES OF ANTICOINCIDENCES (P–E) IN COUNTS/HOUR

thickness of absorber T		Rates (P–E) with absorbing material	
g /cm.²	cm. Pb equivalent	lead	paraffin
0	0	0·051 ± 0·008	0·051 ± 0·008
11·3	1	0 129 ± 0·016	0·117 ± 0·013
22 6	2	0 140 ± 0 014	0·148 ± 0·010
56·5	5	0 207 ± 0·027	—
113	10	0·276 ± 0 030	—

The increase observed is statistically significant. Both the increases observed for the 11·3 g /cm.² absorbers amount to more than three times the standard deviations

No difference between the rates observed under equivalent thicknesses of lead or paraffin is indicated. Therefore the results suggest that the transition effect is mass proportional

Moreover, excluding statistical fluctuations exceeding three times the standard, it is found that the ratio of the effects in lead and paraffin must be less than

$$1{\cdot}8 \cdot 1.$$

Thus the cross-sections for the processes can be assumed to depend on Z as

$$Z^{1{\cdot}0 \pm 0{\cdot}2}.$$

It appears likely that the effect is actually mass proportional This result is in good agreement with the observations of Jánossy & Rochester (1944), this point is discussed further below

It is noted that the relatively accurate measurement of the transition effect was only rendered possible by the use of an anticoincidence arrangement Because of the use of anticoincidences the zero rate was much reduced, and thus the effects produced by an absorber T could be measured accurately. Without the use of anticoincidences the time required to measure accurately the transition effect in small thicknesses would have been prohibitively long

The use of anticoincidences (P–E) instead of coincidences (P) has advantages much more important than the mere reduction of the background

Anticoincidences are recorded whenever a shower falls on P which does not discharge the tray E Note, however, that the tray E is very much larger than the tray t Furthermore, at least three particles are needed to discharge the small tray t, while one particle is sufficient to discharge the large tray E. Thus a shower of reasonably uniform density will hardly discharge the trays t_1, t_2, t_3 without affecting the tray E It must be concluded therefore that the anticoincidences (P–E) are almost exclusively due to showers which are *not extensive*. The term *local penetrating showers* is introduced for the showers which discharge P but not E

In the presence of an absorber T it is clear that the local penetrating showers are showers which are produced in the absorber T by primaries falling on T.

This process was discussed in detail previously, and it can be assumed safely that the local showers are groups of mesons produced by fast nucleons of range about 100 g./cm.^{2}.

The anticoincidences recorded with $T = 0$ need, however, special consideration. As the tray E is so very much more sensitive than the tray t it is immediately clear that even in the absence of T the anticoincidences $(P–E)$ must be due to local showers rather than extensive showers. The question arises as to the point of origin of these local showers. It appears very unlikely that the local penetrating showers in the absence of T could be produced in the air. Because of the low density of air and the long range of the primaries these showers are unlikely to be produced very near to the apparatus, and if they are produced several metres away, they will have only a slight chance of hitting the trays t, etc., in the right places But if some of the local showers contain sufficiently numerous particles to discharge P these showers will have a still larger chance of discharging the large tray E nearby. Therefore local showers coming from the air above will be more likely to produce coincidences (P, E) than anticoincidences $(P–E)$

It was suggested by Dr J G. Wilson that the anticoincidences $(P–E)$ with $T = 0$ might be due to local penetrating showers produced in the lead *below* the tray t, the tray t being discharged by particles emitted upwards This suggestion is much supported by the fact that photographs of penetrating showers quite frequently show particles ejected in the backward direction

It is reasonable to conclude therefore that the anticoincidences $(P–E)$ with $T = 0$ are due to this process, since a similar transition for ordinary cascade showers is of course well known

3·2. *Geometry of the transition effect*

Observations were carried out with an absorber R (roof) 1·5 m above the tray t. The absorber R covered an area of about 2×2 m.^{2}, and it consisted of 8 cm of lead (figure 3)

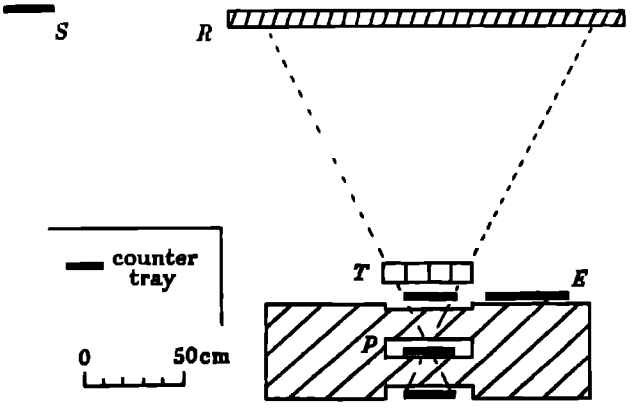


FIGURE 3. The arrangement with roof of 8 cm. lead.

On the basis of the process outlined above, the local penetrating showers are expected to be affected as follows. The thickness of R is of the order of the mean range of the primary nucleons, and so only about one-half or one-third of the primaries are expected to penetrate R without suffering absorption The transition effect observed in the absorber T below the roof is expected to be diminished, then, by this same factor. Primary nucleons absorbed in the roof give rise, of course, to penetrating showers there, and these will contribute a large background to the transition effect unless they can be eliminated This can be done in two ways (1) because of the distance between the set P and the roof it is expected that some at least of the showers generated in R will be too diffuse on reaching P to be recorded, (2) even if some of the showers are dense enough to be recorded it is likely that these will also discharge the large tray E placed 50 cm. from P below the roof To indicate an extensive shower falling on the arrangement a tray of counters S of 1000 cm^2 in area was placed near R at the same level as R itself Anticoincidences (P–E–S) were recorded, these are coincidences (P) unaccompanied by discharges in either E or S, and from what has been said above it is expected that these will be largely due to showers generated in T by those primaries which pass through R without being absorbed.

The rates of anticoincidences (P–E–S) for various thicknesses of T are collected in table 2 below. For the sake of comparison we give in the last line of table 2 the rates given in table 1 for the anticoincidences without R.

Table 2. Anticoincidences (P–E–S) observed under $R = 8$ cm Pb in counts/hour

	absorber T		
event	0	28 cm. paraffin 22·6 g./cm.^{2}	10 cm. Pb 113 g./cm.^{2}
(P–E–S) with R	0·086 ± 0·014	0·123 ± 0·019	0·200 ± 0·023
(P–E) without R	0·051 ± 0·008	0·148 ± 0·010	0·276 ± 0·030

Note that the rate of anticoincidences with $R = 8$ cm Pb and $T = 0$ is considerably smaller than the rate to be expected for $T = 8$ cm Pb and $R = 0$ This difference demonstrates clearly the spreading out of the showers produced in the roof 1·5 m. above the apparatus

The results of table 2 are expressed graphically in figure 4, the two curves being reduced to the same rate for $T = 0$

The increase in anticoincidence rate from $T = 0$ to $T = 10$ cm. Pb is observed to be smaller for $R = 8$ cm. Pb than for $R = 0$, but the difference between the two amounts to only 2·7 times the standard deviation, and therefore it cannot be taken as completely significant Unfortunately, owing to technical difficulties, the series could not be sufficiently extended to establish this difference with certainty, and accordingly it is not possible, as was hoped, to give an accurate estimate of the

range of the primaries in terms of the thickness of R and the decrease in transition effect observed below it The difference, however, if real, does correspond to a range of the order 100 g /cm.^{2}.

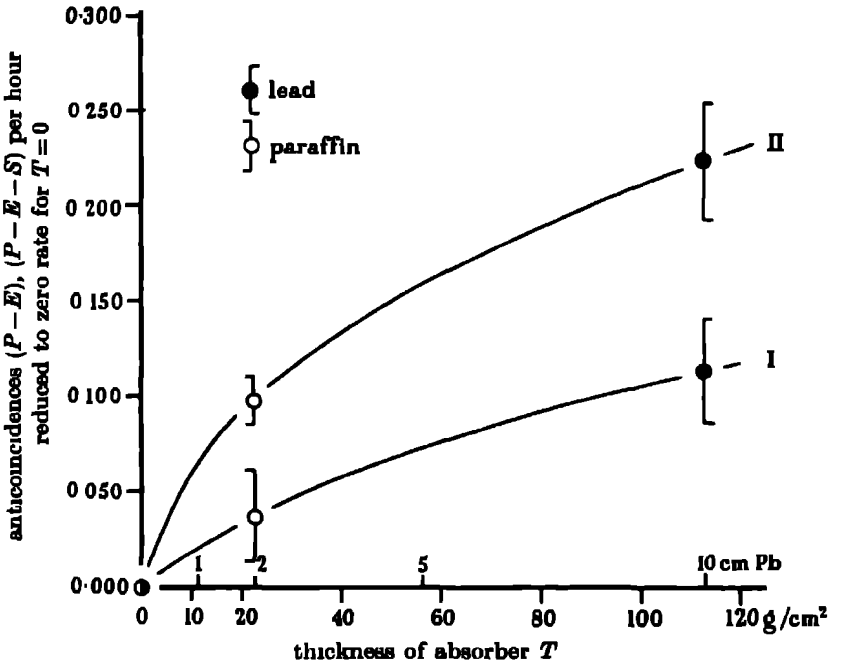


FIGURE 4. Transition effect below roof of 8 cm. lead.
Curve I, with roof, curve II, without roof

3 3. *Transition effect of extensive penetrating showers*

Coincidences (P, E) are caused by extensive penetrating showers. We measured the transition effects of the extensive penetrating showers by recording the rate of coincidences (P, E) as a function of the thickness of the absorber T.

The rates of coincidences thus obtained are collected in table 3, and shown graphically in figure 5.

TABLE 3 RATES OF COINCIDENCES (P, E) IN COUNTS/HOUR

thickness of absorber T		Ratio (P, E) with absorbing material	
g./cm.^{2}	cm. Pb equivalent	lead	paraffin
0	0	0 108 ± 0·012	0·108 ± 0·012
11·3	1	0·180 ± 0·019	0 104 ± 0·012
22·6	2	0 182 ± 0 016	0·102 ± 0·009
56·5	5	0·164 ± 0 024	—
113	10	0 088 ± 0 017	—

It will be seen from table 3 that the rate of coincidences increases with increasing lead absorber T up to a certain thickness The rate decreases again, and for $T = 10$ cm. Pb the rate is much the same as for $T = 0$. The transition effect is reminiscent of the Rossi transition effect of cascade showers While the lead transition of the extensive showers is well established, no transition effect is observed when paraffin absorbers are used, the difference between the rates for lead and paraffin being statistically significant. Thus the transition effect of the extensive showers appears strongly Z-dependent.

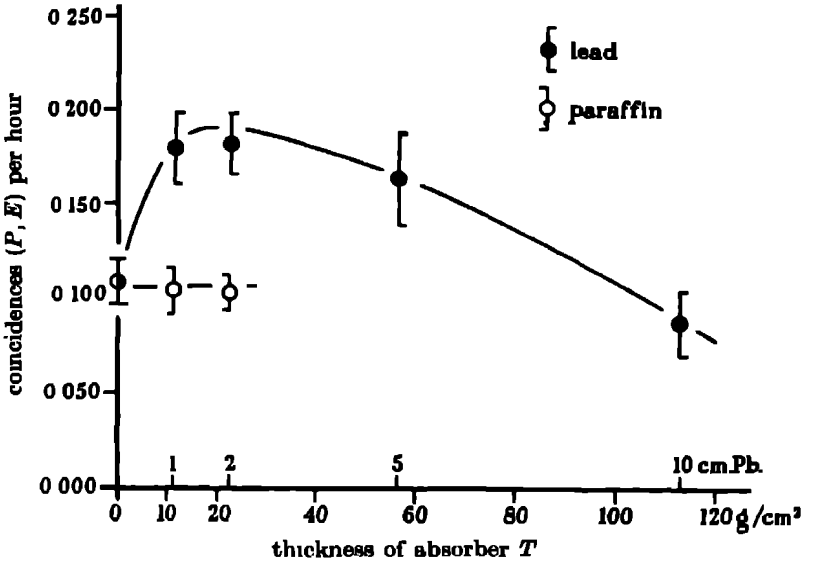


FIGURE 5. Transition effect of extensive penetrating showers.

The transition effect of the extensive air showers can be clearly understood in terms of electron cascades. Assuming that the penetrating extensive showers are mainly cascades the following effects would be expected The cascade falling on a moderately thick lead absorber will increase in density owing to the production of secondary cascades in the lead. The shower, after having increased its density, will be in a more favourable position for setting off the tray t than it was before. Therefore more extensive showers will be recorded in the presence of the lead absorber T than without such an absorber.

If the absorber T is too thick, say 10 cm. of Pb, then it will absorb the cascades rather than increase their density, and thus the rate of recorded coincidences is expected to decline again for such thicknesses.

The details of the transition effect will be discussed in a following paper, but we make these remarks here to clarify the fact that the transition effect declines again after its first rise.

For $T = 0$ showers are recorded which contain three electrons discharging t and also give rise in the absorbers about the apparatus to penetrating particles dis-

charging m and b For $T = 2$ cm. Pb each electron falling on it will give rise to a cascade, and this one electron will be sufficient to discharge t. Consequently the rate is larger with $T = 2$ cm. Pb than with $T = 0$. For $T = 10$ cm. Pb the cascades formed in T will be absorbed and all trays t, m, and b have to be discharged by penetrating particles. This causes a more stringent selection of showers than there was for $T = 2$ cm Pb and consequently the shower rate falls off again.

In this discussion we leave open the question of the origin of the penetrating particles which set off the trays m and b

Regarding the paraffin results, it may be noted that 28 cm of paraffin represent a small fraction of one cascade unit, and therefore no noticeable effect of the paraffin absorber on the density of the cascade is to be expected. Thus the experimental result that no transition takes place in paraffin can be understood immediately.

4 Comparison with previous observations

(i) Note that the transition effect of penetrating showers as observed with *coincidences* (P) *only* represents the superposition of two effects, namely, the transition effect of local showers and that of the extensive showers. Jánossy & Rochester (1944) noticed that the transition effect of lead is better represented by the superposition of two exponential functions, one corresponding to a mean free path of 1 cm. of Pb, the other to a mean free path of 4·6 cm. of Pb, than it is represented by the function of the latter component only. Our present results confirm fully this suggestion, the two components given in 1944 agreeing well with the effects due to extensive and local showers

(ii) Further, the discrepancy between older measurements of the Z-dependence of the transition effect may possibly be cleared up at this stage. Observing with coincidences the transition effect should show a certain Z-dependence, the magnitude of the effect being determined by the relative sensitivity of the recorder to local and extensive showers

The strong Z-dependence found in the earlier work (Jánossy 1942) might have been due to the fact that the t tray of the arrangement used in the 1942 experiments was much less selective to local penetrating showers than the t tray of the arrangement used in the present experiments Therefore the earlier arrangement was relatively more sensitive to the extensive showers, and it is conceivable that the difference in results was due to this.

These points are emphasized in figure 6, which shows the transition effect obtained from the coincidences (P) only. Each curve is, of course, the sum of the corresponding curves of figures 2 and 5. The Z-dependence due to the admixture of extensive showers is very clearly shown.

(iii) Moreover, if the curve of Jánossy & Rochester (1944) were to be corrected to eliminate the transition effect of extensive penetrating showers, the resulting curve would suggest a mean range for the primaries rather higher than that actually obtained and quoted above. This corrected range would tend to minimize the

difference which these authors (1943) found in the ranges of the ionizing and non-ionizing primaries (protons and neutrons) responsible for the local penetrating showers. This brings the results into better agreement with theory, which predicts equal ranges for both protons and neutrons.

(iv) Note, too, that the experiments of the above authors (1943) upon the non-ionizing primaries will be affected because of the back-emission process suggested by Dr Wilson; some non-ionizing primaries will not be observed as such owing to the backwards radiation discharging the anticoincidence counters. Accordingly, the ratio of non-ionizing to ionizing primaries may be much nearer unity than it appeared from previous work

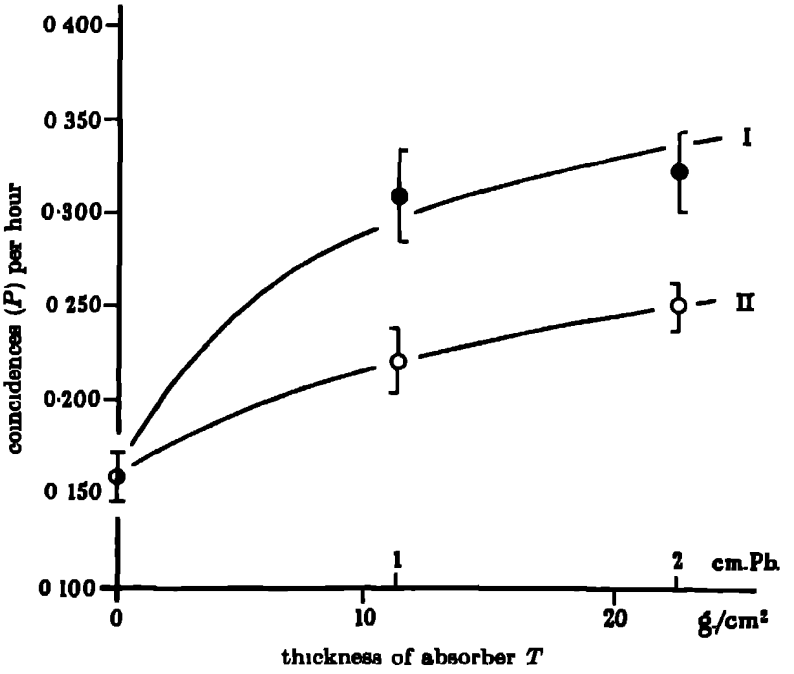


FIGURE 6 Transition effect of coincidences (P).
Curve I, lead, curve II, paraffin.

5. CONCLUSIONS

From the experimental data presented we conclude that the penetrating showers can be split into two independent groups.

(1) Local penetrating showers which are produced in absorbers near the apparatus by single primaries. These showers are most likely due to fast nucleons giving rise to groups of mesons (and probably recoil nucleons).

(2) Extensive air showers which are large cascades containing penetrating particles.

The main argument in favour of the two groups representing different types of showers is that the two types of showers show different transition effects The local penetrating showers are produced at a rate proportional to mass in different materials. The extensive showers, on the other hand, show a strong transition effect in lead but no observable effect in paraffin.

Furthermore, as will be examined in greater detail in a subsequent paper, the rate of coincidences (P, E) is very little changed when the tray E is moved between distances of 9 and 0·5 m. This suggests that the extensive showers observed are all parts of very large showers and makes it unlikely that the purely local showers are due to the same process

Finally, the local penetrating showers contain only moderate numbers of particles, while the extensive showers are spread over large areas and contain great numbers of particles. Details of the investigation of the extensive penetrating showers will be found in a subsequent paper.

The present investigation was started in collaboration with Dr G D Rochester, and we are greatly indebted to him for help and many discussions

References

Jánossy, L. 1942 *Proc. Roy. Soc.* A, **179**, 361.
Jánossy, L. & Rochester, G. D. 1943 *Proc. Roy. Soc.* A, **182**, 180.
Jánossy, L & Rochester, G D. 1944 *Proc. Roy. Soc.* A, **183**, 181.

The crystal structure of some molecular complexes of 4:4′-dinitrodiphenyl

II. The complexes with benzidine, *NN*:*N′N′*-tetramethylbenzidine, 4:4′-dihydroxydiphenyl, 4-iododiphenyl and 4-bromodiphenyl

By D H. Saunder, M.Sc.

University of Cape Town (Beit Railway Trust Rhodesian Fellow)

(*Communicated by Sir Lawrence Bragg, F R S.—Received* 14 *December* 1946)

[Plate 22]

Approximate structures are described for the complexes of 4:4′-dinitrodiphenyl with benzidine, *NN·N′N′*-tetramethylbenzidine, 4:4′-dihydroxydiphenyl, 4 iododiphenyl and 4-bromodiphenyl. A comparison of the more important spectra, of the diffuse reflexions due to thermal vibrations, of the cell dimensions and of photographs taken about structurally equivalent directions, leads to the conclusion that these structures are all of the same general type, and are closely similar to the structure of the complex between 4 4′-dinitrodiphenyl and 4-hydroxydiphenyl, described in detail in part I

Introduction

Several complexes of 4 4′-dinitrodiphenyl with 4- and 4·4′-substituted diphenyl compounds have been investigated and, although they vary considerably in composition and in crystal symmetry, the general type of structure is undoubtedly the same for all of them. As it was not practicable to undertake Fourier analyses for all these complexes the structure of only one of them, the complex between 4:4′-dinitrodiphenyl and 4-hydroxydiphenyl (A), described in part I (Saunder 1946), was determined in detail, and approximate structures were worked out for the complexes of 4.4′-dinitrodiphenyl with *NN N′N′*-tetramethylbenzidine (B), benzidine (C), 4:4′-dihydroxydiphenyl (D), 4-iododiphenyl (E), and 4-bromodiphenyl (F), which are described in the present paper, where, for brevity, these complexes will frequently be referred to by the letters in brackets Of particular interest are the complexes with 4-iodo- and 4-bromodiphenyl, for which effects apparently analogous to optical ghosts due to periodic errors in the structures have been observed, these effects are not discussed in detail in this paper but are dealt with in a later one (James & Saunder 1947)

Comparative X-ray diffraction photographs have also been taken for several other complexes in this series, mentioned in a previous paper (Rapson, Saunder & Stewart 1947). These photographs all show the same characteristics as corresponding photographs for the complexes described here.

A general idea of the type of structure to which all these complexes approximate may be given as follows The dinitrodiphenyl molecules pack in face-centred array as shown, for example, in figure 1 (*d*), and each molecule shown in the diagram represents several such molecules lying one above another, parallel to one another and separated by 3·7 A. In a continuous arrangement of this type long tubular cavities are left running through the structure, and the other component molecules of the complex, which are afterwards often referred to in this paper as the characteristic molecules, occupy these cavities and are seen end-on in the figure The ratio of the number of dinitrodiphenyl molecules to the number of characteristic molecules in the complex thus depends on the length of the characteristic molecules, rather than on chemical considerations

Crystallography

The complexes were all prepared and analyzed in a similar manner, which is described elsewhere (Rapson *et al* 1947) Suitable crystals were grown by slowly cooling a solution of the complex in acetone containing a slight excess of the more soluble component The crystallographic details are summarized in table 1

Table 1

		complex					
		(A)	(B)	(C)	(D)	(E)	(F)
ratio of dinitrodiphenyl to characteristic component		3:1	4.1	4 1	3.1	7.2	7 2
colour		yellow	dark red	red	orange-yellow	pale yellow	cream
system and most prominent faces		monoclinic {100} {110} {001}	triclinic {100} {010} {001}	monoclinic {100} {110} {001} {101}	monoclinic {100} {110}	monoclinic {100} {110}	monoclinic {100} {110
cell angles		$\beta = 90° 39'$	$\alpha = 127° 0'$ $\beta = 103° 39'$ $\gamma = 95° 45'$	$\beta = 120° 28'$	$\beta = 95°$	$\beta = 100°$	$\beta = 100°$
optic axial angle, $2V$		45°	—	. .	29°	34°	37°
and principal refractive indices	α	1·59	—		1 62	1 62	1 60
	β	1 64	—	—	1 63	1 65	1 64
	γ	2 03	—	—	1·99	2 13	2·39
cell dimensions in A	a	20 06	30 2	19 1	20 0	20·0	20 0
	b	9 46	11 15	14 8	18 65	9 5	9 5
	c	11 13	11 5	22 0	11 3	25 8	c 25 8
density in g /c c.		1 43	1 43	1 44	1 45	1 56	1 52
calculated number of complex groups in the unit cell		1 99	2 02	4 05	3 98	2 04	c. 2 07

The directions of the principal refractive indices for the complexes (D), (E) and (F) in the unit cells are very similar to those for the complex (A), shown in figure 1 of part I, with the exception that α is parallel to and γ perpendicular to c in each case. The monoclinic angles, β, were determined from Weissenberg photographs taken with the crystals rotating about the symmetrical axes.

The oscillation photographs for the complexes (E) and (F) are abnormal, and a large number of the spots on these photographs are diffuse and elongated in the *b*-direction instead of being sharply defined These diffuse spectra are possibly the result of periodic errors in the structure, of a nature discussed in detail in another paper (James & Saunder 1947) In oscillation photographs about the *c*-axis for these two complexes only the 0th, 7th and 14th layer-lines consist of sharply defined spectra, all the other layer-lines consist of diffuse spectra. The layer-lines are evenly spaced for the iodo-complex (E), but are not quite regularly spaced for the bromo-complex (F). Only an approximate value can therefore be assigned to the *c*-spacing for the complex (F), and there is, of course, no true unit cell for this compound. In this paper, however, this slight irregularity has been disregarded, and the diffuse spectra have been treated as normal spectra and indexed in the usual manner.

For the complex (B) spectra of all types occur and the space-group may be $P1$ or $P\bar{1}$ For the complex (C) spectra *hkl* occur only with $h+k$ even, and *h*0*l* occur only with both *h* and *l* even; the space-groups C_c and $C_{2/c}$ are therefore possible, but it will be shown from packing considerations that C_c is the more likely space-group For both the complexes (E) and (F) spectra *hkl* occur only with $h+k$ even, *h*0*l* only with *h* even and 0*k*0 only with *k* even, the space-group may therefore be C_m, C_2 or $C_{2/m}$, but it will be shown that C_m is the more likely space-group All these crystals (C), (E) and (F) were so small that attempts to distinguish between the space-group possibilities by testing for the piezo-electric effect were unsatisfactory. With the complex (D), however, the space-group is fixed uniquely as $P_{2_1/a}$, since spectra *hkl* of all types occur, while *h*0*l* occur only with *h* even, and 0*k*0 only with *k* even.

The structures

A comparison of the unit-cell dimensions suggested that all the complexes described in this paper have structures very similar to that of the complex (A), which was described in part I. Assuming this similarity and taking into account the symmetry elements of the space-group it was found that the molecules could be packed into the unit cells so as to give reasonable structures These approximate structures were then tested by observation of the strength of the more important spectra and by observation of the diffuse reflexions due to thermal vibrations Further confirmation was then obtained by taking photographs for the various complexes about *structurally* equivalent directions, these photographs show very striking similarity

In this paper the approximate structures will first be described and then compared with the structure of the complex (A).

The complex with tetramethylbenzidine (B)

The space-group may be either P_1 or $P_{\bar{1}}$, and the unit cell contains two of the complex groups $[O_2NC_6H_4C_6H_4NO_2]_4$ $[(CH_3)_2NC_6H_4C_6H_4N(CH_3)_2]$. The approximate structure described here neither specifically contains nor excludes a centre of symmetry; it need not be materially altered to fit either space-group. Figure 1 (*a*)

shows a projection of the unit cell along a on to the bc plane. The tetramethylbenzidine molecules, seen end-on, are shown lying along a and there are two such molecules, one above the other, in the distance a. Since there is no halving of the unit cell there must, however, be some difference in orientation between successive tetramethylbenzidine molecules in the a-direction. These successive molecules, therefore, are probably inclined differently to the a-axis, they must, however, still lie very nearly in the (011) plane in order to account for the very great strength of the spectrum 011.

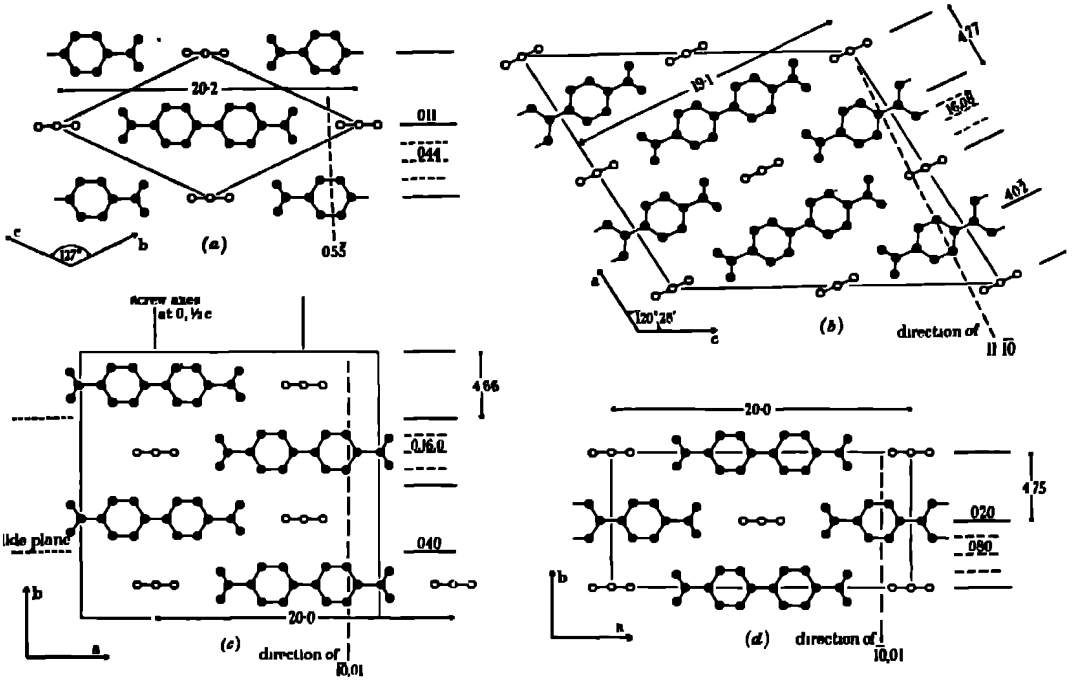


FIGURE 1 (*a*) Projection of the structure of the complex (B) along a on to the bc plane In the distance a there are two tetramethylbenzidine molecules, seen end-on. Each dinitrodiphenyl molecule represents eight molecules separated by $\frac{1}{8}a$. (*b*) Projection of the structure of the complex (C) along b on to the ac plane. The benzidine molecules are seen end-on. Each dinitrodiphenyl molecule represents four molecules separated by $\frac{1}{4}b$ (*c*) Projection of the structure of the complex (D) along c on to the ab plane. The dihydroxydiphenyl molecules are seen end-on Each dinitrodiphenyl molecule represents three molecules separated by $\frac{1}{3}c$. (*d*) Projection of the structure of the complexes (E) or (F) along c on to the ab plane In the distance c there are two iodo- or bromodiphenyl molecules, seen end-on Each dinitrodiphenyl molecule represents seven molecules separated by $\frac{1}{7}c$.

Each dinitrodiphenyl molecule shown on the projection (figure 1 (*a*)) represents eight molecules lying parallel to one another and separated by $\frac{1}{8}a$ or 3 78 A Of course, if the tetramethylbenzidine molecules are inclined to the a-axis, then the eight dinitrodiphenyl molecules will not lie exactly above one another as shown. In order to account for the great strength of the 011 spectrum they must, however, all lie with the long axes of the molecules very nearly in the (011) plane as shown

If the planes of the benzene rings of these molecules are very nearly perpendicular to the (011) plane the dinitrodiphenyl molecules will interleave with one another as shown in figure 2 (*a*), which is a projection of the structure along the axis [01$\bar{1}$] on to the plane (01$\bar{1}$) In this projection the dinitrodiphenyl molecules, lying across the (011) planes, are seen end-on, and the tetramethylbenzidine molecules are seen sideways-on. It will be seen that the dinitrodiphenyl molecules lie in the ($\bar{1}\bar{6}$,11) planes, and, as would be expected from such an arrangement, the spectrum $\bar{1}\bar{6}$,11 and the spectra 800 and $\bar{8}$11 are found to be very strong.

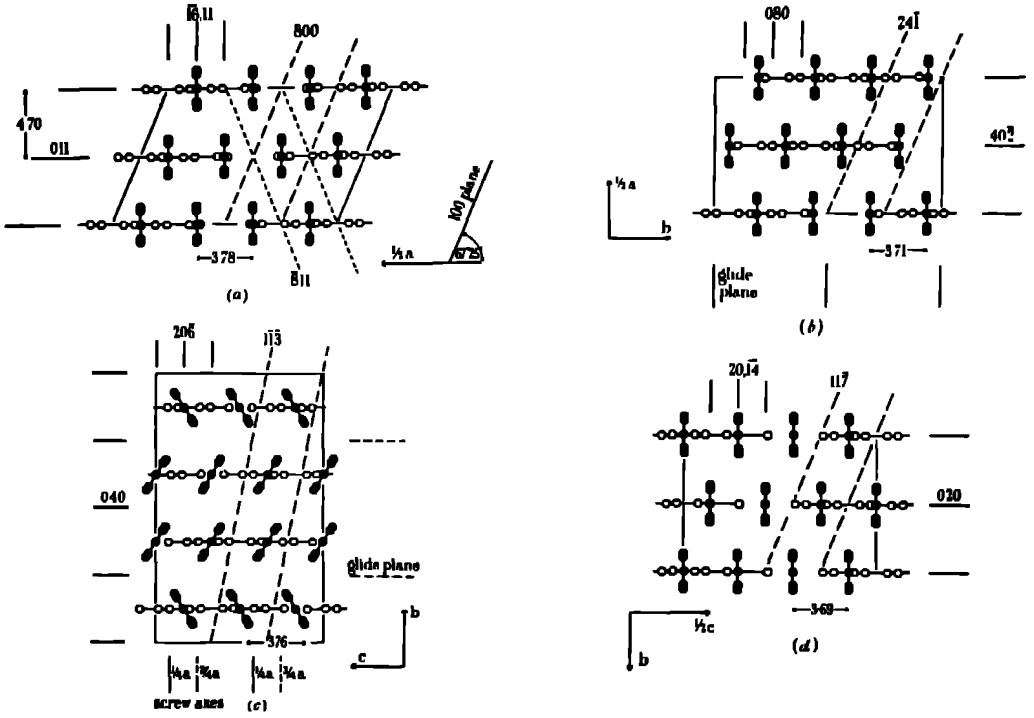


FIGURE 2 (*a*) Projection of the structure of the complex (B) along [00$\bar{1}$] on to the (01$\bar{1}$) plane. The dinitrodiphenyl molecules are seen end-on, and the tetramethylbenzidine molecules sideways-on. (*b*) Projection of the structure of the complex (C) along [102] on to the *ab* plane. The dinitrodiphenyl molecules are seen end-on, and the benzidine molecules sideways-on (*c*) Projection of the structure of the complex (D) along [301] on to the *bc* plane The dinitrodiphenyl molecules are seen end-on, and the dihydroxydiphenyl molecules sideways-on (*d*) Projection of the structure of the complexes (E) or (F) along [701] on to the *bc* plane. The dinitrodiphenyl molecules are seen end-on, and the iodo- or bromodiphenyl molecules sideways-on.

The complex with benzidine (C)

The space-group is probably C_c, as will be shown, and the unit cell contains four of the complex groups $[O_2NC_6H_4C_6H_4NO_2]_4$ $[H_2NC_6H_4C_6H_4NH_2]$ Figure 1 (*b*) shows a projection of the unit cell, which is *c* face-centred, along *b* on to the *ac* plane, which is also a glide plane with translation $\frac{1}{2}c$. The benzidine molecules lie along *b* and are thus seen end-on. Each dinitrodiphenyl molecule shown on the projection

represents four molecules lying parallel to one another and separated by $\frac{1}{4}b$ or 3·71 A. These dinitrodiphenyl molecules must all lie with their long axes very nearly in the $(40\bar{2})$ planes, while the benzidine molecules must lie almost wholly in these planes; this is necessary in order to account for the very great strength of the $40\bar{2}$ spectrum.

It will be seen from figure 2 (*b*), which is a projection of the structure along the axis [102] on to the *ab* plane, that the dinitrodiphenyl molecules lie parallel to the *ac* plane and are contained by the planes midway between the planes (080), this is necessary in order to account for the great strength observed for the spectrum 080, as well as for the spectra $24\bar{1}$ and $24\bar{1}$ With this arrangement, which is consistent with the space-group C_c, it will be observed (figure 2 (*b*)) that the dinitrodiphenyl molecules in adjacent $(40\bar{2})$ planes interleave with one another If the space-group were $C_{2/c}$ the twofold axis parallel to *b* would make it necessary for adjacent dinitrodiphenyl molecules in adjacent $(40\bar{2})$ planes to have the same *b* co-ordinates, and no interleaving of the molecules is then possible. This would mean a closer than normal approach of the molecules to one another, and the spectrum 040 should then be strong, whereas it is actually very weak indeed, as would be expected from the arrangement shown in figure 2 (*b*).

The complex with dihydroxydiphenyl (D)

The space-group is $P_{2_1/a}$, and the unit cell contains four of the complex groups $[O_2NC_6H_4C_6H_4NO_2]_3\,[OHC_6H_4C_6H_4OH]$. Figure 1 (*c*) shows a projection of the unit cell along *c* on to the *ab* plane, the symmetry elements are included in the diagram The special positions given to the molecules are the only ones which satisfy the symmetry conditions and at the same time give a structure closely similar to that of the complex (A), shown in figure 7 of part I. The dihydroxydiphenyl molecules seen end-on are shown lying directly along *c*, but the length of this axis (11 3 A) is not sufficient for this, since the length of a dihydroxydiphenyl molecule, allowing for the approach of the next molecule and even assuming hydrogen bonding between adjacent OH groups, is approximately 12·8 A. These molecules must, therefore, be inclined to *c*, but, in order to account for the very great strength of the 040 spectrum, they must lie almost wholly in the planes midway between the (040) planes, as shown.

Each dinitrodiphenyl molecule shown in figure 1 (*c*) represents three molecules lying parallel to one another and separated by $\frac{1}{3}c$ or 3 76 A. Since the dihydroxydiphenyl molecules are inclined to *c*, the dinitrodiphenyl molecules cannot lie, as shown, directly one above the other; in order to account for the great strength of the spectrum 040, however, they must lie with their long axes very nearly midway between the (040) planes These molecules do not lie parallel to the *ab* plane, but, as for the complex (A), are probably inclined to it so that their lengths lie in the $(20\bar{6})$ planes This is shown in figure 2 (*c*), which is a projection of the structure along the axis [301] on to the *bc* plane, the dinitrodiphenyl molecules are then seen end-on and the dihydroxydiphenyl molecules sideways-on. The *c* co-ordinates of the dinitrodiphenyl molecules were chosen in the only possible way to account for the

great strength of the spectra $11\bar{3}$ and $1\bar{1}\bar{3}$; it was then found that in order to get interleaving of the molecules, and consequently better packing, it was necessary that they should be inclined to the $(20\bar{6})$ planes as shown in figure 2 (*c*). The spectrum $20\bar{6}$ should then be very much weaker for the complex (D) than it was for the complex (A), where the dinitrodiphenyl molecules are practically contained by the $(20\bar{6})$ planes, and this was found to be the case.

The complexes with iododiphenyl (E) *and bromodiphenyl* (F)

Except for the irregularity in the spacing of the layer-lines about the *c*-axis for the bromo-complex, the dimensions of the two crystal cells are the same within the errors of measurement The unit cells contain two of the complex groups $[O_2NC_6H_4C_6H_4NO_2]_7$ $[IC_6H_4C_6H_5]_2$ or $[O_2NC_6H_4C_6H_4NO_2]_7$ $[BrC_6H_4C_6H_5]_2$, and the space-group may be C_m, C_2 or $C_{2/m}$ The face-centring of the structure already accounts for the presence of two complex groups in the unit cell, if there were to be a twofold axis as well, the molecules would have to lie in special positions in order to avoid a further doubling. We need consider only the polar bromo- or iodo-molecules. As these molecules lie with their lengths nearly along *c* and with two molecules in the distance *c*, they must be so arranged that the one molecule is derived from its neighbour by operation of the twofold axis, which must be parallel to *b*. This means that the molecules must lie so that the bromine or iodine atoms are alternately adjacent to and opposed to one another, pairs of bromine or iodine atoms are then spaced along *c* at intervals of approximately 25·8 A instead of being spaced singly at intervals of 12·7 to 12·9 A as they would be if the molecules lay end to end. Now, the arrangement of the layer-lines on rotation photographs for the bromo-complex about the *c*-axis is such as would appear if a set of planes of spacing 3·69 A (probably corresponding to the dinitrodiphenyl molecules) had been modified by some periodic distribution (probably corresponding to the bromodiphenyl molecules) repeating itself in a distance 12 7 A in the direction of the *c*-axis A similar arrangement involving a modifying distribution with a repeat of 12·9 A, will explain corresponding photographs for the iodo-complex This question is fully discussed in the following paper by James & Saunder (1947) on p. 518. It appears likely, therefore, that in both these complexes the bromo- or iododiphenyl molecules lie end to end approximately along *c*; the possibility of a twofold axis is then excluded, and the space-group is probably C_m, which is the same space-group as that possessed by the complex (A).

Figure 1 (*d*) shows a projection of the structure of these complexes along *c* on to the *ab* plane The iodo- or bromodiphenyl molecules lying along *c* are seen end-on, and there are two molecules in the distance *c*. Each dinitrodiphenyl molecule represents seven molecules lying parallel to one another and separated by $\frac{1}{7}c$ or 3·69 A. If the space-group is C_m, in order to avoid doubling the number of complex groups in the unit cell, it is necessary for the iodo- or bromodiphenyl molecules to lie exactly in the mirror planes (020) and for the dinitrodiphenyl molecules to lie across these mirror planes as shown. This was also the case for the complex (A) and

FIGURE 3

(*Facing p. 511*)

is supported by the very great strength of the spectrum 020. It will be seen from figure 2 (*d*), which is a projection of the unit cell along the axis [701] on to the *bc* plane, that the dinitrodiphenyl molecules lie practically in the planes $(20,\overline{14})$, this is necessary in order to account for the great strength of the spectra $20,\overline{14}$, $11\overline{7}$ and $1\overline{17}$

General comparison and confirmation of the structures

The structures of the complexes described here are all of the same general type as that of the complex (A) This is particularly apparent if figure 7 of part I is compared with figure 1 (*a*) to (*d*) in this paper These are projections of the structures of the complexes (A), (B), (C), (D) and (E) or (F) along the axes *c*, *a*, *b*, *c* and *c* respectively Comparative photographs taken with the different crystals mounted about these axes show very great similarity indeed, Weissenberg photographs, showing only the zero layer-line, and taken with the crystals rotating about these axes, are reproduced in figure 3, plate 22. Further photographs taken with the crystals rotating about other structurally equivalent axes also show considerable similarity

In all these structures the dinitrodiphenyl molecules lie *across* one of the crystal planes while the characteristic molecules lie in this plane with their lengths nearly normal to the planes of the benzene rings of the dinitrodiphenyl molecules These crystal planes are, for the complexes (A), (B), (C), (D), and (E) or (F) respectively, the planes (020), (011), $(40\overline{2})$, the planes midway between (040), and the planes (020) (see figure 7 of part I and figure 1 (*a*) to (*d*) in this paper) These planes contain, therefore, a very large number of atoms, namely, four carbon and two nitrogen atoms for each dinitrodiphenyl molecule, and *all* the atoms in the characteristic molecule It is found that in every case the spectrum arising from these planes is exceedingly strong, and is the strongest spectrum appearing in each photograph of figure 3, plate 22 The fourth order spectrum for these planes, namely, the spectrum 080, 044, $16,0\overline{8}$, 0,16,0 and 080 for the complexes (A), (B), (C), (D), and (E) or (F) respectively, is also found to be very strong indeed, this is to be expected, since *all* the atoms of the dinitrodiphenyl molecules now have a nearly maximum contribution (see figure 1 (*a*) to (*d*))

The separation of the equivalent planes (020), (011), $(40\overline{2})$, (040) and (020), in the complexes (A), (B), (C), (D), and (E) or (F) respectively, should be very similar and is found to be always about 4·7 A (see table 2) The distance between the characteristic molecules, lying in these planes, which are separated from one another by dinitrodiphenyl molecules should also be similar and is found to be always about 20 A (again see table 2)

In all these complexes the characteristic molecules lie approximately along one of the principal axes, namely the *c*, *a*, *b*, *c* and *c* axes, for the complexes (A), (B), (C), (D), and (E) or (F) respectively and are nearly normal to all the dinitrodiphenyl molecules, which are separated by approximately 3·7 A in this direction The length of this axis must, therefore, be sufficient to accommodate the right number of dinitrodiphenyl molecules and should be comparable with the length of the characteristic molecule This is shown to be the case in table 2.

TABLE 2

(Dimensions in A)

	complex				
	(A)	(B)	(C)	(D)	(E) or (F)
separation (d) of planes containing characteristic molecules	d(020) 4·73	d(011) 4 70	d(40$\bar{2}$) 4 77	d(040) 4·66	d(020) 4 75
separation of characteristic molecules in above planes	20 1	20·2	19 1	20 0	20 0
length of axis along which characteristic molecules lie	c 11·13	a 30·2	b 14·8	c 11·3	c 25·8
approximate length of characteristic molecule	11 7	15·2	14·0	12·8	12 9 (iodo) 12 7 (bromo)
separation of dinitrodiphenyl molecules in direction of above axis	3·71	3 78	3 71	3·76	3 69

As with the complex (A), described in part I, pronounced diffuse reflexions, due to thermal vibrations, have been observed for all these complexes and provide a useful check on the proposed structures The strongest *circular* diffuse reflexions are found to accompany the spectra 800 and $\bar{8}11$, $24\bar{1}$ and $24\bar{1}$, $11\bar{3}$ and $1\bar{1}\bar{3}$, and $11\bar{7}$ and $1\bar{1}\bar{7}$ for the complexes (B), (C), (D), and (E) or (F) respectively. The directions of the planes corresponding to these spectra have been drawn in on figure 2 (a) to (d), and will be seen to be nearly parallel to the planes of the benzene rings of the dinitrodiphenyl molecules. The strongest diffuse *streaks* are found to accompany the spectra $05\bar{5}$, $11,\bar{1}0$, $\bar{1}0,01$ and $\bar{1}0,01$ for the complexes (B), (C), (D), and (E) or (F) respectively. The direction of the planes producing these spectra have been drawn in on figure 1 (a) to (d) and will be seen to be nearly perpendicular to the lengths of the dinitrodiphenyl molecules

Further confirmation of the relation between the structures of the complexes (A), (D), (E), and (F) is given by the similarity both in direction and magnitude, of the principal refractive indices (see table 1)

THE BONDING

In the detailed analysis of the complex (A), described in part I, it was shown that all the molecules are approximately equally spaced from one another and the closest intermolecular distances of approach are about 3 2 to 3 6 A These distances are of the same order as those observed in crystals of ordinary aromatic nitro-compounds, bonding in which is generally attributed to ordinary van der Waals forces.

In the complexes described in this paper there is no necessity to assume intermolecular approaches closer than those observed for the complex (A); there is sufficient space in the unit cell in each case for the molecules to pack quite normally This is reflected in the densities of the crystals, which are all of the same order as that for the complex (A)

In all these complexes, there is therefore no evidence for localized bonding between the units in the crystal structure; and the molecular ratios in which the

components unite seems to be determined solely by geometrical considerations, the type of structure being such that the number of dinitrodiphenyl molecules that can be accommodated for each characteristic molecule depends on the length of this molecule. In a previous paper (Rapson *et al* 1947), where also the question of bonding is more fully discussed, this condition has been shown to hold for several other complexes in this series.

In conclusion I wish to thank Professor R W. James and Dr W S. Rapson for their constant interest and advice throughout the course of this work, and Mr E. Theal Stewart for the preparation and analysis of most of the complexes. To the Beit Railway Trust Board of Southern Rhodesia I am indebted for a research fellowship.

References

James, R. W & Saunder, D H 1947 *Proc. Roy. Soc* A, 190, 518.
Rapson, W S , Saunder, D. H & Stewart, E. T. 1947 *J. Chem. Soc.* (in the Press)
Saunder, D. H. 1946 *Proc. Roy. Soc.* A, 188, 31.

Description of Plate 22

Figure 3. Zero layer-line Weissenberg photographs for: (*a*) The complex (A) rotating about the *c*-axis, (*b*) the complex (B) rotating about the *a*-axis, (*c*) the complex (C) rotating about the *b*-axis, (*d*) the complex (D) rotating about the *c*-axis, (*e*) the complex (E) rotating about the *c*-axis, (*f*) the complex (F) rotating about the *c*-axis.

Some apparent periodic errors in the crystal lattice of the molecular complexes of 4:4′-dinitrodiphenyl with 4-iodo- and 4-bromodiphenyl

By R W James and D. H Saunder

The University of Cape Town

(*Communicated by Sir Lawrence Bragg, F R S.—Received* 14 *December* 1946)

[Plate 23]

Oscillation photographs from crystals of the complexes of 4 4′ dinitrodiphenyl with 4-bromo- and 4 iododiphenyl show a mixture of sharp and diffuse spots When the oscillation is about the *c*-axis the spots on every seventh layer line are observed to be sharp, those on all the other lines being diffuse The diffuseness varies from line to line, but in all cases corresponds to an elongation of the maxima of the interference function in the reciprocal-lattice space in the *b**-direction The spacing of the diffuse layer-lines in the c-axis photographs from the complex with 4-bromodiphenyl is anomalous, and the positions of the lines correspond to those of optical ghosts accompanying the sharp layer-lines, such as would be expected if a fundamental spacing of 3 60 A were periodically distorted, the period of the distortion being rather less than 3½ times that of the fundamental spacing The diffuseness of the spots suggests that the structure also contains faults in the *b*-direction A suggested explanation of the effects in terms of the structure is put forward

In the second part of the paper, a mathematical discussion of diffraction by a simplified model, consisting of a lattice with a periodic error in spacing in the c-direction, and faults consisting of sudden random changes in the phase of a periodic error in lattice planes perpendicular to the *b**-axis, is discussed, and it is shown that the diffraction pattern given by such a structure exhibits features similar to those observed.

Introduction

The crystal structures of a number of molecular complexes of 4 · 4′-dinitrodiphenyl with other 4- and 4 4′-substituted diphenyl compounds have been investigated by one of us (Saunder 1946, 1947) Although the compounds examined vary considerably both in crystal symmetry and in the ratio of the two components of the complex, the underlying structure is undoubtedly of the same general type for all of them. The typical arrangement may be taken as that in the complex of 4 4′-dinitrodiphenyl with 4-hydroxydiphenyl in the ratio of three molecules of the first to one of the second, the structure of which has been fully determined (Saunder 1946). Two of the complexes examined, those with 4-bromo- and 4-iododiphenyl, give X-ray photographs showing certain peculiarities, which are of considerable interest from the optical point of view, and which it is the purpose of this paper to discuss

A general idea of the type of structure to which all complexes of the kind considered approximate may be given as follows The dinitrodiphenyl molecules form layers in face-centred array, as shown in figure 1. In any given crystal, the molecules

will not in general lie exactly in these idealized planes, but they never depart greatly from them Suppose a series of such planes to lie one above the other with a spacing of about 3·7 A. The dinitrodiphenyl molecules in any one layer are denoted by *A* in figure 1, and those in any row parallel to *a* are so placed that a gap is left between adjacent nitro groups When a series of superposed layers is considered, it will be seen that these gaps produce a set of tubular cavities, also in face-centred array, running through the structure. The characteristic molecules of the complex, in the cases to be considered, the 4-bromo- and 4-iododiphenyl molecules, occupy these tubular cavities They are seen end-on in figure 1, and are denoted by *B*. The general

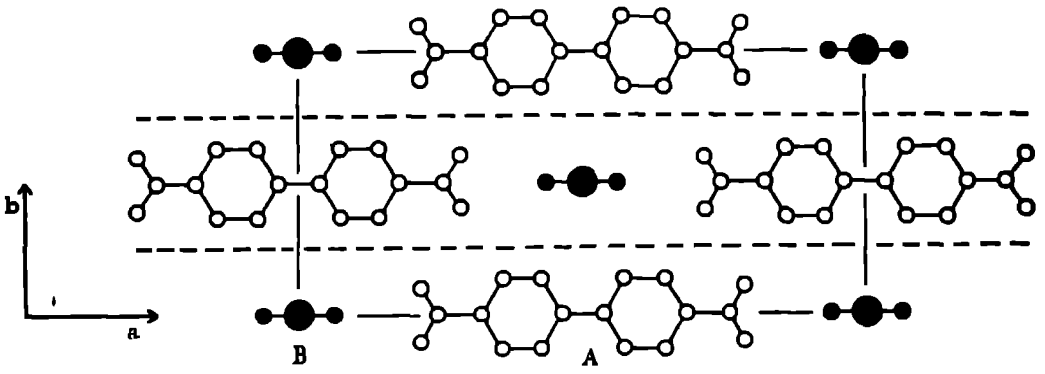


FIGURE 1. Projection along *c* on to the *ab*-plane of the idealized structure of the complexes of 4 4′-dinitrodiphenyl with 4 iodo- and 4-bromodiphenyl In the distance c there are two iodo- or bromodiphenyl molecules (*B*), seen end-on Each dinitrodiphenyl molecule (*A*) represents seven molecules one above the other, separated by *c*/7.

direction of the long axes of these *B* molecules is perpendicular to the planes containing the *A* molecules. In the only structure fully worked out, mirror planes of symmetry lie perpendicular to the planes of *A* molecules, and contain their long axes and also the planes of the benzene rings of the *B* molecules This exact arrangement is not found in all the complexes, although it is probably also that of the iodo- and bromo-complexes discussed here, but the departure from it never seems to be large

The ratio of the number of *A* to the number of *B* molecules varies from complex to complex, but appears to depend on the length of the *B* molecules, rather than on chemical considerations, and, indeed, it has been found possible to predict the ratio from a knowledge of this length (Rapson, Saunder & Stewart 1947). In general, the *A* molecules do not lie exactly in planes, nor is the direction of the length of the *B* molecules perpendicular to that of the lengths of the *A* molecules, and on account of these and other departures from the ideal structure outlined above, the actual symmetry of the crystals may differ widely from case to case, but consideration of the actual structures shows the departure from the ideal type to be relatively slight, and oscillation photographs taken about structurally equivalent axes are extraordinarily similar for all the complexes.

The complexes with 4-iodo- and 4-bromodiphenyl

The preparation and chemical analysis of these complexes has been described by Rapson *et al.* (1947) The molecular ratio of dinitrodiphenyl to iodo- and bromodiphenyl is found to be 7 · 2 in each case, so that the complexes may be represented by the formulae

$$(IC_6H_4C_6H_5)_2(O_2NC_6H_4C_6H_4NO_2)_7 \quad \text{and} \quad (BrC_6H_4C_6H_5)_2(O_2NC_6H_4C_6H_4NO_2)_7$$

Details of the crystallographic determinations will be found in a previous paper (Saunder 1947) The crystal symmetry is monoclinic. As will be seen later, the arrangement of the spots in oscillation photographs from the bromo-complex is in some respects abnormal, but with the iodo-complex the positions of the spots on the photographs are normal and show the dimensions of the unit cell to be

$$a = 20\,0\,\text{A}, \quad b = 9\,5\,\text{A}, \quad c = 25\,8,\text{A} \quad \beta = 100°.$$

The space-group may be C_m, C_2, or $C_{2/m}$, but there is reason to think that of these C_m is the most likely, which is the same as that for the complex with 4-hydroxydiphenyl, the structure of which has been fully determined. Detailed comparison of the intensities of the more important spectra given by this compound with those given by the iodo-complex, and comparison of the cell dimensions, the refractive indices, and the diffuse spectra due to thermal vibrations, leads to the conclusion that the structures of the two crystals must be extremely similar

If the abnormalities in the spectra from the bromo-complex are for the moment disregarded, all the remarks given above may be taken as applying to it also, the dimensions of the two unit cells being the same within the limits of error of the photographs taken.

Discussion of the oscillation photographs

In the crystals of these complexes, the planes most closely corresponding with the idealized planes containing the *A* molecules in figure 1 are the (001) planes, and if the planes of *A* molecules were regularly spaced the distance between successive planes would be 3·69 A If the structure consisted only of such regularly spaced planes of *A* molecules, a rotation photograph taken about the *c*-axis would show widely spaced layer-lines corresponding to this distance In the actual complex with iododiphenyl, the pattern repeats itself every seven layers of *A* molecules, there being two *B* molecules in this distance, so that intermediate layer-lines make their appearance dividing the distance between the layer-lines due to the *A* molecules into seven equal parts, thinking in terms of the reciprocal-lattice space. Actual oscillation photographs taken about the *c*-axis show the layer-lines corresponding to this larger spacing, 7 × 3 69 A; their appearance is, however, abnormal The 0th, 7th and 14th layer-lines, corresponding to a spacing of 3 69 A, which on an ideal structure would be given by the layers of *A* molecules alone, are sharp and clear, but all the intermediate layer-lines show spots that are more or less diffuse, with

elongations along the layer-lines. The diffuseness of the spots on any one line appears to be about the same, but it varies markedly from line to line There is no trace of diffuseness in the *A* layer-lines, which are in every way normal. A photograph from the bromo-complex is reproduced in figure 2, plate 23 Alternations of sharp and diffuse layer-lines have also been observed by Powell & Huse (1943) from molecular complexes of picryl chloride with hexamethylbenzene, and from several other complexes of a similar type. They have ascribed these to faults in the regular sequence of atomic planes in structure

Oscillation photographs about the *b*-axis show sharp and diffuse spots also (see figure 4, plate 23), and in every case the diffuse spots are those that would lie on the intermediate layer-lines in photographs about the *c*-axis On the photographs taken with rotations about the *b*-axis the diffuse spots are elongated transversely to the layer-lines, as would, of course, be expected from the *c*-axis photographs, but the elongated spots are not always centred on the layer-lines, but show marked displacements above and below them The degree of this displacement appears to be a property of the individual crystal, since it varies in photographs obtained from different specimens

A careful study of photographs taken with oscillations about different axes shows the elongation of the maxima of the interference function about the corresponding points in the reciprocal-lattice space to be parallel to b^*. An elongation in this direction would be produced by a crystal having only a small extension in the *b* direction, but this would make all the spots equally diffuse, which is not observed here The explanation of the diffuseness is therefore probably to be sought in faults of the nature of those discussed by Wilson (1942), which are superposed on the framework of a fundamentally regular lattice structure. A possible type of fault which might account for the effect will be considered later

Abnormalities observed in photographs from the complex with 4-bromodiphenyl

Photographs taken with oscillations about the *c*-axis with the bromo- and iodo-complexes are at first sight very similar (figures 2 and 3, plate 23, respectively). Both show the sharp 0th, 7th and 14th layer-lines, and the diffuse intermediate ones. Careful inspection shows, however, that in the photograph from the bromo-complex the spacing of the intermediate layer-lines is not uniform, when considered in terms of the reciprocal-lattice space, but that the lines are arranged in groups, themselves evenly spaced, about the sharp layer-lines. So far as the position of the lines alone is concerned, the arrangement is exactly that to be expected of a series of optical ghosts accompanying the sharp layer-lines, such as would appear if a set of planes of regular spacing 3·69 A had been modified by a periodic error of spacing repeating itself in a distance 12·7 A in the direction of the *c*-axis, that is to say, with two repetitions in a distance a little less than that corresponding to seven layers of the *A* molecules Had the error repeated itself *exactly* twice in seven layers, the ghost

layer-lines would have exactly divided the spacing between the zero layer-line and the 7th layer-line, and between the 7th and 14th layer-lines, as in fact they do within the errors of measurement in the case of the iodo-complex

On the actual photographs from the bromo-complex the departure from regularity is very plain to the eye The measured spacings of the layer-lines in the reciprocal-lattice space, as determined by means of the Bernal chart, and using Cu K_α radiation, are shown diagrammatically in figure 5 The notation used in the figure has the following significance Let n be the order of a spectrum from a series of regularly spaced planes. Suppose these planes to be displaced periodically, a complete period of the displacement occurring every Q planes Then additional spectra of order $n \pm m/Q$ make their appearance, and these are the optical ghosts. Any spectrum is denoted by the symbol (n, m), the spectra from the undisturbed planes being those of order $(n, 0)$ When the planes are periodically displaced, the spectra of order $(n, 0)$ still appear, but those with $m \neq 0$ also appear. In the photographs from the bromo-complex, the positions of the layer-lines correspond to a value of Q rather less than $3\frac{1}{2}$ †

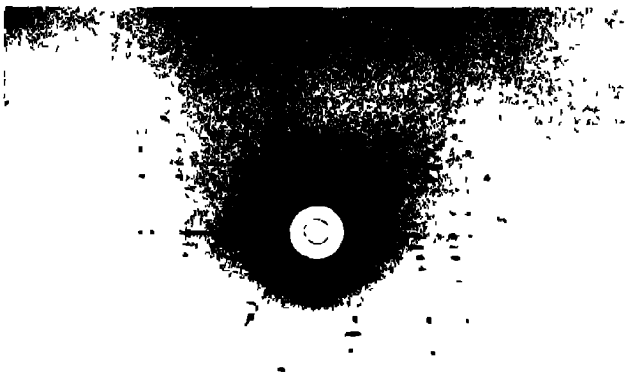
The *positions* of the irregularly spaced layer-lines are explained in detail by the assumption of a periodic error in spacing According to this view, the layer-lines, in the order in which they occur in the photograph, are $(0, 0)$, $(1, \bar{3})$, $(0, 1)$, $(1, \bar{2})$, $(0, 2)$, $(1, \bar{1})$, $(0, 3)$, $(1, 0)$, $(0, 4)$, $(1, 1)$, $(2, \bar{2})$ and $(0, 5)$, $(1, 2)$, $(2, \bar{1})$, $(1, 3)$, $(2, 0)$, . As a working hypothesis, the occurrence of the B molecules at regular intervals might be supposed to impose a variation of spacing on the sheets of A molecules It is significant that the error period in the case of the bromo-complex, as determined by direct measurement from the photographs, is 12 7 A, which is the length, as closely as can be determined from the available data, to be expected for the molecule of 4-bromodiphenyl, and twice this length, which is 25 4 A, is rather less than $7 \times 3{\cdot}69$, which is 25 8 A. The length to be expected for the molecule of 4-iododiphenyl is 12·0 A, and twice this is equal to 25 8 A The periodic error then becomes part of the regularly repeating pattern of the crystal, and the ghost layer-lines become ordinary layer-lines

On the view that the intermediate layer-lines are ghost spectra due to a periodic error in a fundamental spacing, the ghosts accompanying the zero layer-line, the spectra $(0, \pm m)$, would be expected to be very weak, although not in principle absent, while in the actual photograph these spectra are quite strong. It must, however, be remembered that the structure contains the regularly spaced B mole-

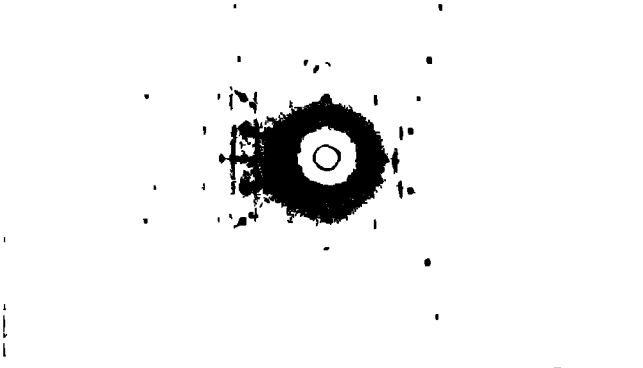
† After this paper had been sent for publication photographs were taken from the complex of 4 4′-dinitrodiphenyl with 4-chlorodiphenyl The irregularity of the spacing of the layer-lines is even more marked than in the case of the complex with the bromo-compound, as had in fact been anticipated. The lines again form groups about the zero and seventh layer-lines, and are consistent with an error period of 12 3 A and ratio of the dinitro constituent to the chloro constituent of 3 3 1 Photographs from the corresponding fluoro-compound are perfectly normal, all spots being sharp, the ratio of the constituents being now 3·1. The photographs resemble extremely closely those from the complex with 4:4′-hydroxydiphenyl. It is hoped to publish further details of these cases in a subsequent paper.



2



3



4

(*Facing p.* 532)

cules, which, considered alone, would give a set of spectra in the positions of the spectra $(0, \pm m)$ so that in any case these would be expected to occur.

Another point requiring consideration is the diffuseness of the lines. The existence of diffuse ghost spectra must be explained, the diffuseness of which, moreover, varies with the order m of the ghost. It is plain that the crystals of these complexes are not simple enough for an exact quantitative explanation of the observed effects to be possible. All that can be attempted with much hope of success is to produce a simplified model that will give spectra exhibiting the same general type of properties as those actually observed In the second half of the paper an attempt will be made to do this.

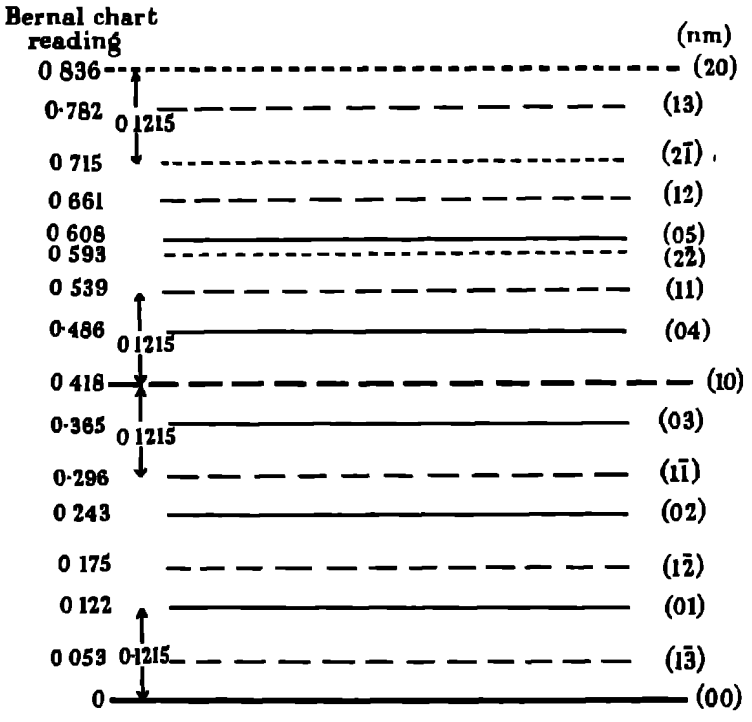


FIGURE 5 Scale representation of the layer-lines as they appear on a rotation diagram for a rotation about the c-axis with the complex with 4-bromodiphenyl

THEORETICAL THE DISCUSSION OF AN IDEALIZED MODEL

Assume the crystal to consist of layers of A molecules, all exactly alike, placed one above the other in the array represented in figure 1, and this array will be assumed to persist coherently throughout the crystal Through this structure the chains of B molecules pass transversely to the A planes, and parallel to the c-axis of the crystal Suppose the B molecules lying in any one chain to maintain an unchanged regular sequence through the crystal. The dotted lines in figure 1 are the traces of planes dividing the crystal into a series of slabs parallel to the planes

(010) In any one such slab the structure is supposed to be without faults, corresponding atoms in the different B chains all having the same c co-ordinates, so that the consequent periodic displacements of the A planes are all in phase. Suppose, however, that in different slabs there are relative displacements of the B molecules, so that in passing from one slab to the next there may be a sudden change in the c co-ordinates of corresponding B molecules It will be assumed that there is a certain probability that within any distance measured parallel to b such a sudden change, or fault, may occur Making the simplest possible assumption, it is seen that there are two possible sets of co-ordinates for the B molecules. It will then be necessary to work out in the manner described by Wilson (1942) the probability that two slabs separated by a given distance shall have the same or different c co-ordinates, and thence to calculate the interference function for the crystal, remembering at the same time that the B molecules introduce periodic variations in the spacing of the A molecules, which will vary in phase according to the occurrence of faults in the slabs Consideration is given first to the scattering due to the A molecules alone, with the periodic errors imposed on their spacing by the B molecules, neglecting for the time being the scattering by the B molecules themselves For this purpose, the model will be still further simplified, each unit of the structure being replaced by a scattering point

Scattering by a periodically distorted point-lattice with faults

Consider a point-lattice that is parallelepipedal in form, having N_1, N_2 and N_3 points parallel to a, b and c respectively. The position relative to the origin of a point in the lattice when it is undistorted is given by the vector

$$\mathbf{r} = u\mathbf{a} + v\mathbf{b} + w\mathbf{c}, \tag{1}$$

u, v, w being integers.

Suppose a periodic error of the c spacing to be introduced, which for the sake of simplicity will be taken as sinusoidal in character, the c co-ordinate of a point being given by $[w + q\cos\{2\pi cg(w+\phi_v)\}]\mathbf{c}$ Here, g is the *frequency* of the periodic disturbance, the number of times it occurs in unit distance, so that $cg = 1/Q$, Q being the number of c spacings within which a complete period of the error occurs The maximum amplitude of the displacement, expressed as a fraction of c, is q, and ϕ_v is a phase constant, which depends on v, and has the same value so long as v is constant, that is to say, over any (010) plane. According to this simple assumption, it will later be supposed that ϕ_v may have the values 0 or ϕ but no other value.

Consider first scattering by a single b-plane of points, with v constant and $\phi_v = \phi$ The vector position of a point on such a plane may now be written

$$\mathbf{r}_v = u\mathbf{a} + v\mathbf{b} + [w + q\cos\{2\pi cg(w+\phi)\}]\mathbf{c}. \tag{2}$$

Let $\mathbf{s}_0$ and $\mathbf{s}$ be unit vectors in the directions of incidence and scattering, and let

$$\mathbf{S} = \mathbf{s} - \mathbf{s}_0, \quad |\mathbf{S}| = 2\sin\theta, \tag{3}$$

2θ being the angle of scattering. One may then write for the amplitude scattered by a single plane v, in terms of that scattered by a single unit,

$$A_v = \sum_{u=0}^{u=N_1-1} \sum_{w=0}^{w=N_2-1} \exp\{2\pi i(\mathbf{S}.\mathbf{r}_v)/\lambda\}. \tag{4}$$

The scattering in terms of a distribution in the reciprocal-lattice space may now be expressed in the usual way, writing

$$\mathbf{S}/\lambda = \xi\mathbf{a}^* + \eta\mathbf{b}^* + \zeta\mathbf{c}^*, \tag{5}$$

which gives, using (2) and the properties of reciprocal vectors,

$$A_v = \exp(2\pi i v\eta) \sum_u \exp(2\pi i u\xi) \sum_w \exp 2\pi i[w\zeta + q\zeta\cos\{2\pi cg(w+\phi)\}]. \tag{6}$$

A_v is now expressed as a function of ξ, η and ζ, and it gives the amplitude scattered in the direction s when the conditions of incidence and scattering are such that the extremity of the vector $\mathbf{S}/\lambda$ lies at the point $(\xi a^*, \eta b^*, \zeta c^*)$ in the reciprocal-lattice space, that is to say, when the sphere of reflexion passes through this point.

Carrying out the summation with respect to u, (6) may be written in the form

$$A_v = \exp(2\pi i v\eta)\frac{\exp(2\pi i N_1\xi)-1}{\exp(2\pi i\xi)-1} \sum_{w=0}^{w=N_2-1} \exp 2\pi i[w\zeta + q\zeta\cos\{2\pi cg(w+\phi)\}]. \tag{7}$$

The summation with respect to w may be handled by a method employed by Kochendorfer (1939) in a paper on the effect of periodic errors in lattices. Now use an expansion in terms of Bessel functions,

$$\exp(ipx) = J_0(p) + 2\sum_{m=1}^{m=\infty} i^m J_m(p)\cos mx, \tag{8}$$

from which follows immediately

$$\exp[2\pi i q\zeta\cos\{2\pi cg(w+\phi)\}] = \sum_{m=-\infty}^{m=\infty} i^{|m|} J_{|m|}(2\pi q\zeta)\exp(2\pi i mcgw)\exp(2\pi i mcg\phi), \tag{9}$$

a result that follows at once if the cosines are expressed in the exponential form.

The sum with respect to w in (7) can now be written

$$\sum_w = \sum_{m=-\infty}^{m=\infty} i^{|m|} J_{|m|}(2\pi q\zeta)\exp(2\pi i mcg\phi) \sum_{w=0}^{w=N_2-1} \exp\{2\pi i w(\zeta + mcg)\}. \tag{10}$$

The sum over w on the right-hand side of (10) is a geometrical progression of a type occurring in the theory of the ordinary diffraction grating, and is equal to

$$\frac{\exp\{2\pi i N_2(\zeta+mcg)\}-1}{\exp\{2\pi i(\zeta+mcg)\}-1} = \exp\{\pi i(N_2-1)(\zeta+mcg)\}\frac{\sin\{\pi N_2(\zeta+mcg)\}}{\sin\{\pi(\zeta+mcg)\}}. \tag{11}$$

The expression (11) has maxima equal to N_2 whenever $\zeta + mcg$ is a whole number, and these maxima are exceedingly sharp if N_2 is large, while, under the same conditions, the subsidiary maxima that occur between the main maxima are negligible.

Considered as a function of ζ therefore, (11) has sharp maxima when $\zeta = l \pm mcg$, l being an integer, and is otherwise of negligible magnitude, a result that will be used later.

Now write (6) in the form

$$A_v = \exp(2\pi i v\eta)\frac{\exp(2\pi i N_1\xi)-1}{\exp(2\pi i\xi)-1}F_v, \tag{12}$$

where $F_v = \sum_w$, and is given by (10), and for the whole crystal the amplitude scattered in the direction considered is

$$A = \sum_v A_v = \frac{\exp(2\pi i N_1\xi)-1}{\exp(2\pi i\xi)-1}\sum_{v=0}^{v=N_3-1} F_v \exp(2\pi i\eta v) \tag{13}$$

The corresponding intensity, expressed in terms of the scattering by a single point, and considered as a function of ξ, η and ζ, is called the interference function for the piece of crystal considered. Its value is

$$I_0(\xi,\eta,\zeta) = |A|^2 = \frac{\sin^2(\pi N_1\xi)}{\sin^2(\pi\xi)}\sum_v\sum_{v'} F_v F_v^* \exp\{2\pi i\eta(v-v')\}, \tag{14}$$

or, if $v' = v+n$,

$$I_0(\xi,\eta,\zeta) = \frac{\sin^2(\pi N_1\xi)}{\sin^2(\pi\xi)}\sum_v\sum_n F_v F_{v+n}^* \exp\{-2\pi i n\eta\}. \tag{15}$$

Now F_v can have one of two possible values, one being that given by (10) and the other the corresponding value with $\phi = 0$. These two values may be called F_ϕ and F_0 respectively Let I_n be the average value of $F_v F_{v+n}^*$ for two b planes separated by n planes, and let M be the number of times this particular separation occurs. Then

$$I_0(\xi,\eta,\zeta) = \frac{\sin^2(\pi N_1\xi)}{\sin^2(\pi\xi)}\sum_n M I_n \exp(-2\pi i n\eta) \tag{16}$$

In the case of a parallelepipedal lattice, $M = N_2 - |n|$, and n can run from $N_2 - 1$ to $-(N_2-1)$ Thus

$$I_0(\xi,\eta,\zeta) = \frac{\sin^2(\pi N_1\xi)}{\sin^2(\pi\xi)}\sum_{n=-(N_2-1)}^{n=N_2-1} I_n(N_2-|n|)\exp(-2\pi i n\eta) \tag{17}$$

The next step is to calculate the average value I_n, and to do this the method given by Wilson (1942, 1943) in dealing with the effects of faults in alloy structures is applied Assume that there are two possible values of F_v, F_0 and F_ϕ, and that a fault may occur at any b plane of such a kind that the value of F_v changes from one possible value to the other Let $P(n)$ be the probability that the plane $v+n$ is like the plane v, and let αds be the probability that a fault occurs between bs and $b(s+ds)$. Consider a plane (2) at a distance $b(s+ds)$ from a plane (1) The plane (2) may be like plane (1) because the plane at a distance bs is like plane (1) and no fault occurs in the further distance bds; or because the plane at a distance bs is unlike plane (1) and a fault occurs in the further distance bds which once more makes plane (2) like

plane (1) The probability that plane (2) is like plane (1) is therefore the sum of the probabilities of these two mutually exclusive types of event. Thus

$$P(s+ds) = P(s)(1-\alpha ds) + \{1-P(s)\}\alpha ds = P(s) - 2\alpha P(s)\,ds + \alpha ds. \tag{18}$$

But since ds is small, then

$$P(s+ds) = P(s) + (dP(s)/ds)\,ds, \tag{19}$$

whence, by comparing (18) and (19),

$$dP(s)/ds = -2\alpha P(s) + \alpha,$$

which, on integration, remembering that $P(0) = 1$, gives

$$P(s) = \tfrac{1}{2}(1 + e^{-2s\alpha}), \tag{20}$$

or, since s gives the distance expressed as a multiple of the lattice spacing b,

$$P(n) = \tfrac{1}{2}\{1 + \exp(-2\,|\,n\,|\,\alpha)\}. \tag{21}$$

The average value of $F_v F^*_{v+n}$, I_n, may be written in the form

$$\begin{aligned} I_n &= \tfrac{1}{2}F_0[P(n)\,F_0^* + \{1-P(n)\}\,F_\phi^*] + \tfrac{1}{2}F_\phi[P(n)\,F_\phi^* + \{1-P(n)\}\,F_0^*] \\ &= \tfrac{1}{2}P(n)\{|\,F_0\,|^2 + |\,F_\phi\,|^2\} + \tfrac{1}{2}\{1-P(n)\}\{F_0F_\phi^* + F_\phi^*F_\phi\} \end{aligned} \tag{22}$$

Now

$$F_0 = \sum_{m=-\infty}^{m=\infty} i^{|m|}J_{|m|}(2\pi q\zeta) \sum_{w=0}^{w=N_3-1} \exp\{2\pi i w(\zeta + mcg)\} = \sum_m B_{0m}(\zeta), \tag{23}$$

$$F_\phi = \sum_{m=-\infty}^{m=\infty} i^{|m|}J_{|m|}(2\pi q\zeta)\exp(2\pi i mcg\phi) \sum_{w=0}^{w=N_3-1} \exp\{2\pi i w(\zeta + mcg)\} = \sum_m B_{\phi m}(\zeta). \tag{24}$$

Each of these expressions is a function of ζ with very narrow maxima, and the only terms in $|\,F_0\,|^2$ and $|\,F_\phi\,|^2$ that are appreciable are those of type $|\,B_{0m}(\zeta)\,|^2$ or $|\,B_{\phi m}(\zeta)\,|^2$, while in $F_0F_\phi^*$ or $F_0^*F_\phi$ the only terms of appreciable value are those of the types $B_{0m}(\zeta)\,B^*_{\phi m}(\zeta)$ or $B^*_{0m}(\zeta)\,B_{\phi m}(\zeta)$ Neglecting all other terms, and using (11), then

$$\tfrac{1}{2}\{|\,F_0\,|^2 + |\,F_\phi\,|^2\} = \sum_{m=-\infty}^{m=\infty} J^2_{|m|}(2\pi q\zeta)\frac{\sin^2\{\pi N_3(\zeta+mcg)\}}{\sin^2\{\pi(\zeta+mcg)\}} = \sum_m S_m, \tag{25}$$

$$\tfrac{1}{2}\{F_0F_\phi^* + F_0^*F_\phi\} = \sum_m S_m\cos(2\pi mcg\phi) \tag{26}$$

By (22), (25), (26) and (21), therefore,

$$\begin{aligned} I_n &= \sum_m [P(n)\,S_m + \{1-P(n)\}\,S_m\cos(2\pi mcg\phi)] \\ &= \tfrac{1}{2}\sum_m [S_m\{1+\cos(2\pi mcg\phi)\} + \exp(-2\,|\,n\,|\,\alpha)\,S_m\{1-\cos(2\pi mcg\phi)\}] \\ &= \sum S_m\{\cos^2(\pi mcg\phi) + \exp(-2\,|\,n\,|\,\alpha)\sin^2(\pi mcg\phi)\} \end{aligned} \tag{27}$$

Equation (17) now becomes

$$I_0(\xi,\eta,\zeta) = \frac{\sin^2(\pi N_1\xi)}{\sin^2(\pi\xi)} \sum_{n=-(N_2-1)}^{n=N_2-1} \sum_{m=-\infty}^{m=\infty}$$

$$\times S_m\{\cos^2(\pi mcg\phi) + \exp(-2\,|\,n\,|\,\alpha)\sin^2(\pi mcg\phi)\}(N_2-|\,n\,|)\exp(-2\pi i n\eta). \quad (28)$$

Now write

$$\sum_{n=-(N_2-1)}^{n=N_2-1}(N_2-|\,n\,|)\exp(-2\pi i n\eta) = N_2 + 2\sum_{n=1}^{n=N_2-1}(N_2-|\,n\,|)\cos 2\pi n\eta \quad (29)$$

$$= \frac{\sin^2(\pi N_2\eta)}{\sin^2(\pi\eta)}; \quad (29a)$$

for (29) is an alternative form of the ordinary grating formula (29a) The right-hand side of (29) is a Fourier series of N_2 terms, and the greater N_2 the sharper are the periodic maxima in η This gives a clue to the nature of the second term in (28), which contains the summation

$$D(\eta) = N_2 + 2\sum_{n=1}^{n=N_2-1}(N_2-|\,n\,|)\exp(-2\,|\,n\,|\,\alpha)\cos 2\pi n\eta. \quad (30)$$

This again is a Fourier series, with maxima for the same values of ζ as those for which (29) is a maximum. The coefficients are, however, multiplied by the factor $\exp(-2\,|\,n\,|\,\alpha)$, which decreases rapidly with increasing $|\,n\,|$ if α is appreciable. The effect of this is to broaden the maxima; for the terms of higher order, to which the sharpening of the maxima is due, become ineffective This sum is therefore denoted by $D(\eta)$ to indicate that its maxima are relatively diffuse

Using (25), (29a) and (30), (28) may now be written in the form

$$I_0(\xi,\eta,\zeta) = \frac{\sin^2(N_1\pi\xi)}{\sin^2(\pi\xi)} \sum_{m=-\infty}^{m=\infty} J^2_{|m|}(2\pi q\zeta)\frac{\sin^2\{\pi N_3(\zeta+mcg)\}}{\sin^2\{\pi(\zeta+mcg)\}}$$

$$\times\left\{\frac{\sin^2(\pi N_2\eta)}{\sin^2(\pi\eta)}\cos^2(\pi mcg\phi) + D(\eta)\sin^2(\pi mcg\phi)\right\}. \quad (31)$$

When $m = 0$,

$$I_0(\xi,\eta,\zeta) = J_0^2(2\pi q\zeta)\frac{\sin^2(\pi N_1\xi)}{\sin^2(\pi\xi)}\frac{\sin^2(\pi N_2\eta)}{\sin^2(\pi\eta)}\frac{\sin^2(\pi N_3\zeta)}{\sin^2(\pi\zeta)}, \quad (32)$$

an expression giving sharp maxima when $\xi = h$, $\eta = k$ and $\zeta = l$, h, k, l being integers. These are in fact the ordinary lattice maxima, which are in the same positions as those given by the undistorted lattice and are just as sharp The *intensities* are, however, modified by the factor $J_0^2(2\pi q\zeta)$, which may be written $J_0^2(2\pi lq)$ for the sharp maxima, and this becomes smaller with increasing l, the rate of decrease depending on q, the maximum distortion.

The ghost spectra

In addition to the main maxima, other maxima occur when

$$\zeta \pm |m|cg = l, \quad \text{or} \quad \zeta = l \pm |m|cg = l \pm |m|/Q,$$

where Q is the number of repetitions of the undisturbed spacing in one repetition of the periodic distortion. These extra maxima are the ghost spectra, a group of which, corresponding to different values of m, for the same value of l accompanies each main maximum The ghost maxima are sharp in ξ and ζ, but their sharpness in η is governed by the factor

$$\frac{\sin^2(N_3\pi\eta)}{\sin^2(\pi\eta)}\cos^2\left(\frac{\pi m\phi}{Q}\right) + D(\eta)\sin^2\left(\frac{\pi m\phi}{Q}\right),$$

and so is a function of m, the order of the ghost, and not of the order of the main spectrum that the ghost accompanies The first term of this factor is sharp if N_3 is large, but the second term is much more diffuse if α, the probability of a fault, is appreciable. The two terms, however, make different relative contributions for different values of m. For small $m\phi/Q$, the first term will be preponderant and the second will have small values, but, as $m\phi/Q$ grows, the importance of the second or diffuse term increases and that of the first or sharp term diminishes For still larger values of $m\phi/Q$ the relative importance of the terms may be once more reversed The form of the factor suggests that the spots on the ghost layer-lines might show sharp nuclei with relatively diffuse wings, and on the photographs from the bromo- and iodo-complexes a number of such spots are in fact to be seen, which may perhaps be some confirmation of the view that the effects observed are due to an irregularity of the same general nature as that discussed

The intensities of the ghost maxima for the simple lattice with a sinusoidal displacement are determined by the factor $J_m^2(2\pi q\zeta)$, which for $m \neq 0$ is small for small values of ζ So long as $2\pi q\zeta$ is not too large, $J_m(2\pi q\zeta)$ is proportional to ζ^m The ghosts of higher order are therefore unimportant for the low orders of main spectra, but become more important for the higher orders, even when the distortion is sinusoidal. The ghosts accompanying the zero-order spectrum will be expected to be faint, since they are given by the values of $J_m^2(2\pi mq/Q)$, which will be small unless the maximum displacement is large The zero-order ghosts are not in principle absent, as the approximate treatment of periodic errors, given, for example, by Daniel & Lipson (1943), seems to indicate.

The assumption of a sinusoidal distortion has been made for the sake of simplicity It is clear that the actual distortion, although periodic, cannot be sinusoidal. It would, in fact, be more accurate to represent it by a Fourier series of the type $c\sum_k q_k \cos(2\pi kcgw)$, instead of by $cq\cos(2\pi cgw)$. If no limitations to small displacements are imposed, the algebra in this case becomes rather heavy. If, however, it is assumed that q_k is always small, it is easy to show that the kth order ghost accompanying any main maximum has an amplitude proportional to q_k, the kth Fourier

coefficient of the series representing the periodic displacement If the displacement is not small, matters are more complicated, and the amplitude of the kth order ghost does not depend on q_k alone, but it appears still to be true that the value of q_k exercises the determining influence

The spectra due to the B molecules

In the actual photographs, the ghosts accompanying the zero layer-line are strong, much stronger than would be expected from the analysis given above It must be remembered, however, that this refers only to the effect of the A molecules of the underlying lattice, distorted, it is supposed, by the B molecules, the scattering by which has so far been neglected In the actual photographs, the scattering by the B molecules is of course recorded The structure must be considered as a whole Before doing so, however, the scattering from the B molecules alone will be considered, assuming the same probability of faults in the structure as was assumed above Consider, as before, the amplitude scattered by the B molecules lying in a single (010) plane, corresponding to a given value of v Write this in the two possible forms

$$A_v = R_1(\xi)\, R_3(\zeta) \exp(2\pi i v\eta),$$

$$A_{v'} = R_1(\xi)\, R_3(\zeta) \exp(2\pi i v'\eta) \exp(2\pi i\phi\zeta), \tag{33}$$

$R_1(\xi)$ and $R_3(\zeta)$ being the amplitude factors obtained by summing over u and w respectively It must be remembered that the frequency in $R_3(\zeta)$ is that corresponding to the repetition of the B molecules, that is to say it is g, the frequency of the periodic error assumed in calculating the effect of the distorted lattice of A molecules Just as before, one writes for the interference function due to the B molecules alone

$$I_B(\xi, \eta, \zeta) = \sum_v \sum_n |R_1(\xi)|^2 |R_3(\zeta)|^2 F_v F^*_{v+n} \exp(-2\pi i n\eta), \tag{34}$$

where F_v and F_{v+n} are now either $+1$ or $\exp(2\pi i\phi\zeta)$

Now introduce the probability of a sudden change from one type of F to the other, just as before This gives for the average value of $F_v F^*_{v+n}$

$$I_n = \cos^2(\pi\phi\zeta) + \exp(-2|n|\alpha) \sin^2(\pi\phi\zeta), \tag{35}$$

and, following exactly the same argument as that given above in deriving (31), then

$$I_B(\xi, \eta, \zeta) = |R_1(\xi)|^2 |R_3(\zeta)|^2 \left\{ \frac{\sin^2(\pi N_2\eta)}{\sin^2(\pi\eta)} \cos^2(\pi\phi\zeta) + D(\eta) \sin^2(\pi\phi\zeta) \right\}. \tag{36}$$

If ζ has the same significance as in the discussion of the scattering by the A molecules, its maxima, and so the layer-lines in a rotation photograph about the c axis, occur when $\zeta = \pm m/Q$, for the c^* spacing of the B reciprocal lattice is $1/Q$ of that of the A lattice. Equation (36) shows therefore that the B molecules alone would produce a set of layer-lines having the same positions, and the same variation of diffuseness with order, as those predicted for the ghost spectra of zero order The

relatively strong spectra actually observed in these positions can therefore mainly be ascribed to this cause, the true ghosts being those accompanying the higher orders

In the case of the iodo-complex, in which the spacing of the sheets of A molecules and the period of the distortion appear to be commensurable, and which is thus to be considered only as a degenerate case of a structure with a periodic error, the higher orders of these $(0, m)$ spectra produced by the B molecules will always coincide with ghost spectra (n, m) accompanying higher orders. In the case of the bromo-complex, in which the two periodicities appear not be to commensurable, this will not be so, and some doubling of the layer-lines might therefore be expected There is some evidence that it does actually occur. Reference to figure 5 will show that the layer lines $(0, 5)$ and $(2, \bar{2})$ lie very close together The layer line $(0, 5)$ is faint, and no $(0, 6)$, or higher orders, are observed at all, but there are one or two spots on the line $(0, 5)$ visible on the actual photographs which quite definitely do not lie on the line $(2, \bar{2})$. Doubling between $(0, 4)$ and $(2, \bar{3})$ is also to be expected, but is not observed in the photographs. It is, however, perhaps significant that this layer-line, and some of the other lines of higher order, are broadened in the c-direction, which might indicate an incipient doubling of the lines Corresponding lines in the photographs from the iodo-complex are not thus broadened

Diffraction by two lattices with incommensurable spacings but with a definite phase relationship

In the above discussion, it has been assumed that the interference function can be built up by adding together the effects of the A and B molecules, considered separately. This, of course, is an over-simplification Interference effects occur that depend on the juxtaposition of the two types of molecule, and these may be very important if the two periodicities are commensurable, so that the two structures, taken together, constitute a single repeating pattern. It is necessary to consider this point in more detail

The simple case of reflexion from two sets of parallel planes, one set with a spacing a_1, the other with a spacing a_2, will now be discussed. Let f_1 and f_2 be the structure factors of a single plane of each set, which may conveniently be supposed as complex, to take into account any phase difference due to the relative positions of the sets of planes Let θ be the glancing angle of reflexion from the planes. Then the phase difference between waves reflected from successive planes of the two sets are respectively μa_1 and μa_2, where $\mu = (4\pi \sin\theta)/\lambda$. If in the piece of crystal considered there are N_1 planes of one set and N_2 planes of the other, the amplitude reflected in the direction θ may be written

$$A = f_1 \sum_{n=0}^{n=N_1-1} \exp(in\mu a_1) + f_2 \sum_{m=0}^{m=N_2-1} \exp(im\mu a_2)$$

$$= f_1 \exp\{\tfrac{1}{2}(N_1-1)\mu a_1\} \frac{\sin(\tfrac{1}{2}\mu N_1 a_1)}{\sin(\tfrac{1}{2}\mu a_1)} + f_2 \exp\{\tfrac{1}{2}(N_2-1)\mu a_2\} \frac{\sin(\tfrac{1}{2}\mu N_2 a_2)}{\sin(\tfrac{1}{2}\mu a_2)}. \quad (37)$$

The intensity I is equal to $|A|^2$, which gives

$$I = |f_1|^2 \frac{\sin^2(\frac{1}{2}\mu N_1 a_1)}{\sin^2(\frac{1}{2}\mu a_1)} + |f_2|^2 \frac{\sin^2(\frac{1}{2}\mu N_2 a_2)}{\sin^2(\frac{1}{2}\mu a_2)}$$
$$+ 2|f_1||f_2| \frac{\sin(\frac{1}{2}\mu N_1 a_1)}{\sin(\frac{1}{2}\mu a_1)} \frac{\sin(\frac{1}{2}\mu N_2 a_2)}{\sin(\frac{1}{2}\mu a_2)} \cos[\tfrac{1}{2}\mu\{(N_1-1)a_1-(N_2-1)a_2\}-\delta], \quad (38)$$

where δ is given by $f_1 f_2^* = |f_1||f_2| e^{i\delta}$, and is the phase difference between waves reflected from the first planes of the two sets If these planes are at a distance d apart, $\delta = \mu d$

The first two terms in (38) give the sum of the intensities produced by the two sets of planes acting independently The effect of their interaction is given by the third term Considered as functions of θ, the two factors involving sines have very sharp maxima when N_1 and N_2 are large, and the value of the last term will therefore be inappreciable unless both factors have principal maxima for the same value of θ If a_1 and a_2 are commensurable, so that $p_1 a_1 = p_2 a_2$, p_1 and p_2 being integers, this will occur for a number of values of θ If p_1 and p_2 have no common factor, the first coincidence will be between the spectrum of order p_2 due to the planes a_1 with that of order p_1 due to the planes a_2, and the next will be between the spectra of order $2p_2$ and $2p_1$, and so on. It is plain that these coincidences produce no spectra in positions other than those already given by the two sets of planes considered independently, but they will modify the intensities of certain of these, because there is a definite phase relationship between the two sets of planes When the spacings are commensurable $(N_1-1)a_1 = (N_2-1)a_2$ in any block of crystal that contains a number of complete repetitions of the pattern, so that the cosine factor in the last term of (38) becomes simply $\cos\delta$

To take a very simple example, the (111) planes of diamond may be considered to consist of two sets of planes with $a_1 = a_2$, one set being displaced relative to the other through a distance $a_1/4$ The spectra due to the two independent sets of planes now coincide for all orders, but since $\delta = 0$ for the zero order, $\pi/2$ for the first, π for the second, $3\pi/2$ for the third, 2π for the fourth, and so on, it will be seen that the intensities of these spectra are respectively $4f^2$, $2f^2$, 0, $2f^2$, $4f^2$, and so on. It is clear that this is merely a slight modification of the ordinary method of calculating the intensity from the structure factor.

If the spacings are not commensurable, coincidences will not occur, and if N_1 and N_2 are large, the last term of (38) will always be negligible, and the spectra observed will be just those due to the two sets of planes considered independently In the case considered in this paper, one set of planes has been assumed, the A planes, to have a periodic *variation* of spacing of the same period as the spacing of the other set, the B planes, and this produces, in addition to the two sets of spectra given by the two sets of planes independently without error of spacing, the sets of ghost spectra accompanying the spectra due to the A planes The effects observed with the bromo-complex appear to be adequately explained by this type of assumption If

the two periodicities are commensurable, as appears to be the case with the iodo-complex, the problem can be treated as a case of a repeating pattern, with a definite structure factor, which can be considered as a degenerate case of a structure with a periodic error In the case actually dealt with, the pattern does not repeat itself exactly at all, so that ordinary methods of structure-factor calculation are not applicable

The treatment of the problem given here must be regarded only as an attempt to produce a model that shall show effects of the same general type as those observed It has not been found possible to account in detail for the patterns obtained from the actual crystals No really satisfactory explanation of the fact that, in the oscillation photographs taken about the *b*-axis, the diffuse spots are not always centred exactly on the layer-lines has been found, although it seems almost certainly to be connected with some imperfection in the crystals on a scale larger than those considered here

Part of the work described in this paper has been carried out during the tenure by one of us (D H S) of a Beit Railway Trust Fellowship.

References

Daniel, V & Lipson, H 1943 *Proc. Roy Soc* A, **181**, 368, **182**, 378.
Kochendörfer, A 1939 *Z Kristallogr* **101**, 149
Powell, H M & Huse, G. 1943 *J. Chem. Soc* (ii), p. 435.
Rapson, W S , Saunder, D H. & Stewart, E T 1947 *J. Chem. Soc* (in the Press)
Saunder, D H. 1946 *Proc. Roy Soc* A, **188**, 31
Saunder, D H 1947 *Proc. Roy. Soc.* A, **190**, 508.
Wilson, A J C 1942 *Proc. Roy Soc.* A, **180**, 277.
Wilson, A. J. C. 1943 *Proc. Roy. Soc.* A, **181**, 360.

Description of Plate 23

Figure 2 The complex with 4-bromodiphenyl 15° rotation about the *c*-axis Rays initially incident along a^*

Figure 3. The complex with 4-iododiphenyl· 15° rotation about the *c*-axis Rays initially incident along a^*

Figure 4 The complex with 4-iododiphenyl 15° rotation about the *b*-axis. Rays initially incident along a^*.

Decay of vorticity in isotropic turbulence

By G. K. Batchelor, *Trinity College, Cambridge*
and A. A. Townsend, *Emmanuel College, Cambridge*

(*Communicated by Sir Geoffrey Taylor, F.R.S.—Received* 17 *December* 1946)

The equation describing the rate of change of the mean square vorticity in homogeneous isotropic turbulence is obtained and the terms occurring therein are discussed. A negative contribution to $d\overline{\omega^2}/dt$ arises from the effect of viscosity, while a positive contribution is produced by the tendency for the random diffusive motion to extend the vortex lines. This latter contribution can be related to the skewness of the probability distribution of the rate of extension of line elements of the fluid aligned in any given direction.

The results of direct measurements of each of the factors appearing in the vorticity equation are then described. The measurements were made by analyzing electrically the output from a hot-wire anemometer placed downstream from a grid in a uniform stream. Both $U^2/\overline{u^2}$ and λ^2 are found to increase approximately linearly with time during decay of the turbulence and their rates of change are consistent with the energy equation. The skewness factor mentioned above is approximately constant during decay, with the same value at all Reynolds numbers. It follows that the rate of increase of $\overline{\omega^2}$ due to vortex extension is proportional to $(\overline{\omega^2})^{\frac{1}{2}}$, and further measurements show that the effect of viscosity has a similar dependence, so that the ratio of the two contributions to $d\overline{\omega^2}/dt$ remains the same throughout the decay. The viscous contribution is always the greater but the contributions tend to equality as the grid Reynolds number increases. The measurements of all terms in the vorticity equation are shown to satisfy the equation with sufficient accuracy. One of the deductions from the measurements is that the double velocity correlation function tends to a cusp at the origin as the Reynolds number increases indefinitely.

1. Introduction. Consequences of isotropy

Considerable progress in the difficult task of understanding the mechanism of turbulent motion in fluids has been made since 1935 when G. I. Taylor introduced the notion of isotropy, or spherical symmetry. Taylor defined homogeneous isotropic turbulence by the property that mean values of all functions of the components of the velocity fluctuation and their derivatives are independent of rotations and reflexions of the axes defining the frame of reference. The mathematical theory simplifies considerably under this assumption and several results of practical significance have been derived. It has also been established that a good approximation to isotropic turbulence can be produced in a wind tunnel by passing the air stream through a mesh of regularly spaced rods, and that velocity fluctuations in the fluid can be accurately recorded with a hot-wire anemometer. Consequently the subject of isotropic turbulence has great interest and offers the possibility of a parallel development of theory and experiment.

The mathematical theory of those aspects of the motion in homogeneous isotropic turbulence which are described by double and triple velocity correlations has been set out by v. Karman & Howarth (1938). These authors showed that the mean values of the products of two and three components of the velocity fluctuations at

two points in the turbulent field could each be specified by a single scalar function of the distance r between the points Thus

$$u'^2 R_i^j = -\frac{f'}{2r}\xi_i\xi_j + (f + \tfrac{1}{2}rf')\,\delta_{ij}, \tag{1·1}$$

$$u'^3 T_{ij}^k = \left(\frac{k - rk'}{2r^3}\right)\xi_i\xi_j\xi_k + \left(\frac{k + \frac{1}{2}rk'}{2r}\right)(\delta_{ik}\xi_j + \delta_{jk}\xi_i) - \frac{k}{2r}\delta_{ij}\xi_k, \tag{1·2}$$

where δ_{ij} is the unit tensor whose value is unity if $i = j$ and zero otherwise. Subscripts to the tensors on the left indicate the directions of velocity components at the point P while superscripts refer to the point P' The vector ξ, of magnitude r, represents the spatial interval PP', and u' is the root mean square of the fluctuation of any velocity component at any point of the field. The functions $f(r)$ and $k(r)$ are the scalars specifying the double and triple velocity correlations respectively (with dashes to denote differentiation with respect to r) and also represent the correlation coefficients for the particular configurations shown in figure 1 f and k/r are even functions of r, the latter vanishing with r The mean square of the component of vorticity in any direction is easily determined from (1·1) with the result

$$\omega'^2 = -5u'^2 f_0'' = \frac{5u'^2}{\lambda^2}, \tag{1 3}$$

where λ is the dissipation length parameter introduced by Taylor (1935)

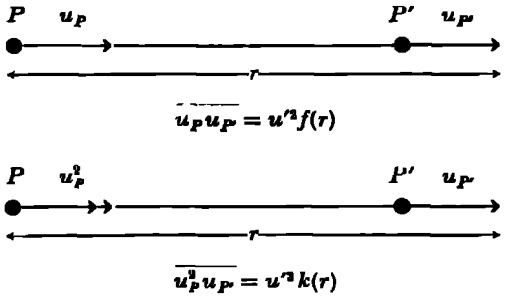


FIGURE 1. Correlation functions $f(r)$ and $k(r)$.

The dynamical aspect of the theory is developed by applying the Navier-Stokes equations at each of the points P and P' An equation for the rate of change of the double velocity correlation with respect to time is thus obtained; this will involve the triple velocity correlation in view of the non-linearity of the equations, but not the pressure since it forms part of a first order tensor which is zero in isotropic turbulence Using (1·1) and (1·2), the dynamical equation becomes a relation between f and k, viz.

$$\frac{\partial(fu'^2)}{\partial t} - u'^3\left(k' + \frac{4}{r}k\right) = 2\nu u'^2\left(f'' + \frac{4}{r}f'\right). \tag{1·4}$$

This is the fundamental equation for the rate of change of the double velocity correlation and expresses all the information concerning the double velocity correlation which can be derived from the Navier-Stokes equations with the single assumption of isotropy.

The equation giving the rate of dissipation of energy of turbulence is obtained from (1·4) by putting $r = 0$, whence

$$\frac{du'^2}{dt} = 10\nu u'^2 f_0'' = -\frac{10\nu u'^2}{\lambda^2} \tag{1·5}$$

Other equations having a physical interpretation may be obtained by comparing the coefficients of other powers of r in equation (1·4). Thus the coefficients of r^2 give

$$\frac{d(\frac{1}{2}f_0'' u'^2)}{dt} - \tfrac{7}{6}u'^3 k_0''' = \tfrac{7}{3}\nu u'^2 f_0^{\mathrm{iv}}, \tag{1·6}$$

and in view of (1·3), this equation describes the rate of change of mean square vorticity.

It is the purpose of this paper to examine this equation for the decay of vorticity in isotropic turbulence. The terms of the equation have each been measured by means of a hot-wire anemometer and suitable electrical analysis of the output. The measurements help to make clear the processes accompanying the decay of turbulence behind a mesh in a wind tunnel.

2 The vorticity equation

The physical significance of the vorticity equation is more easily seen when it is derived directly from the Navier-Stokes equations. As v. Karman (1937) has shown, this direct deduction leads in the case of homogeneous isotropic turbulence to

$$3\frac{d\omega'^2}{dt} = 2\sum_{i,j}\overline{\omega_i\omega_j\frac{\partial u_i}{\partial x_j}} + 2\nu\sum_{i,j}\overline{\omega_i\frac{\partial^2\omega_i}{\partial x_j^2}}, \tag{2·1}$$

where ω_i is the component of vorticity in the i-direction. The second term on the right side clearly represents the entire effect of viscosity on the mean square vorticity, since

$$2\overline{\omega_i\frac{\partial^2\omega_i}{\partial x_j^2}} = \frac{\partial^2\overline{\omega_i^2}}{\partial x_j^2} - 2\overline{\left(\frac{\partial\omega_i}{\partial x_j}\right)^2}, \tag{2·2}$$

and since $\partial^2\overline{\omega_i^2}/\partial r_j^2$ is zero in a homogeneous field, the term is essentially negative and represents a rate of destruction of vorticity. Referring to (1·6) and (1·3), it follows that f_0^{iv} is always positive as is already known from the measured shape of the correlation curve.

The first term on the right side of (2·1) can be written

$$2\sum_{i,j}\overline{\omega_i\omega_j\frac{\partial u_i}{\partial x_j}} = 2\sum_i\overline{\omega_i\,|\,\omega\,|\frac{\partial u_i}{\partial s}} = 2\,\overline{\omega^2\left(\frac{\partial u}{\partial s}\right)_s}, \tag{2·3}$$

where $\mathbf{u}$ is the local total velocity vector, and ω^2 is the square of the local total vorticity vector s represents arc length of the local vortex line, and when written as a subscript denotes a component parallel to that arc length Taylor (1938) discussed the meaning of this term and demonstrated both theoretically and experimentally that it is positive and represents a rate of production of vorticity. Note first that $(\partial \mathbf{u}/\partial s)_s$ is the instantaneous rate of extension of the local vortex line Now when positive extension of a vortex filament occurs, the magnitude of the local vorticity increases due to the consequent lateral contraction and angular acceleration. Thus in parts of the fluid where there is a positive rate of extension of the vortex filament, the magnitude of the vorticity will be high, so that one may expect $\omega^2(\partial \mathbf{u}/\partial s)_s > 0$ Taylor pointed out that this is consistent with the general tendency in turbulent motion for points to move apart and for lines to lengthen, giving a tendency for the average value of ω^2 to increase. He was the first to explain that this production of vorticity due to random, diffusive extension of vortex lines is a fundamental process in the mechanics of turbulence and shows the reason for the very high rate of dissipation of turbulence energy

The effect of extension of the vortex lines is to tend to make the vorticity distribution 'spotty', with small regions of high vorticity, on the other hand, the effect of viscosity is strongest in regions of high vorticity, and tends to diffuse it evenly throughout the fluid Equation (2·1) represents the balance between these two effects and an important deduction from the experimental results to be described hereunder is that this balance is *similar* for the decay range tested. A single indirect measurement of the value of the terms of equation (2 1) has been described previously (Taylor 1938) and the result was that the rate of increase of ω'^2 due to vortex extension and the rate of decrease due to viscous dissipation were, in this particular experiment, respectively about three and four times the resultant rate of decrease of ω'^2

A further useful expression for the rate of increase of ω'^2 due to vortex extension is obtained by deducing from (1·2) that

$$\overline{\left(\frac{\partial u}{\partial x}\right)^3} = u'^3\left(\frac{\partial^3 T_{11}^1}{\partial \xi_1^3}\right)_{r=0} = u'^3 k_0''',$$

where u is the turbulent velocity component in the direction of the x-axis. Now the skewness of a probability distribution which is not too far from a normal distribution is commonly measured by the dimensionless ratio formed from the third and second moments If S is minus the skewness factor of the probability distribution of $\partial u/\partial x$ (the minus sign is introduced because the skewness is found to be negative), then

$$S = -\overline{\left(\frac{\partial u}{\partial x}\right)^3} \Big/ \left[\overline{\left(\frac{\partial u}{\partial x}\right)^2}\right]^{\frac{3}{2}} = -k_0''' \lambda^3 \qquad (2\cdot4)$$

The contribution to $d\omega'^2/dt$ from the process of vortex extension is thus (from 1·6) directly proportional to S, evidently the rate of extension of the vortex lines is

related to the average over space of the mean cube of the rate of extension in a particular direction.

The various mathematical expressions of the vorticity equation can thus be written:

(1) contribution to $d\omega'^2/dt$ due to extension of the vortex lines

$$= \frac{2}{3}\overline{\omega^3\left(\frac{\partial \bar{u}}{\partial s}\right)_s} = -\frac{35}{3}u'^3 k_0''' = -\frac{35}{3}\overline{\left(\frac{\partial u}{\partial x}\right)^3} = \frac{7}{3\sqrt{5}}\omega'^3 S, \tag{2·5}$$

(2) contribution to $d\omega'^2/dt$ due to the effect of viscosity

$$= -\frac{2}{3}\nu\sum_{i,j}\overline{\left(\frac{\partial \omega_i}{\partial x_j}\right)^2} = -\frac{70}{3}\nu u'^2 f_0''^{\mathrm{v}} = -\frac{14}{3\sqrt{5}}\omega'^3\frac{G}{R_\lambda}, \tag{2·6}$$

where

$$R_\lambda = \frac{u'\lambda}{\nu}, \quad G = \lambda^4 f_0''^{\mathrm{v}}$$

The form most suitable for comparison with experiment is

$$\frac{d\omega'^2}{dt} = \frac{7}{3\sqrt{5}}\omega'^3 S - \frac{14}{3\sqrt{5}}\omega'^3\frac{G}{R_\lambda}. \tag{2·7}$$

Each of the terms in equation (2·7) has been measured using methods which are described in the following section.

3. Experimental method

The wind tunnel used in the experiments is in the Cavendish Laboratory, Cambridge, and was erected by Professor G. I. Taylor and Dr D. C. MacPhail. It is a closed return tunnel designed to provide a low level of turbulence in the working section.* The working section is initially 15 × 15 in. and is 72 in long. A slight divergence of the walls provides nearly zero pressure gradient over the whole section at commonly used speeds. In order to protect the anemometer wires, care has been taken to seal the tunnel from the outside air except at the slot used to insert the meshes. Turbulent flow which is approximately isotropic is produced by introducing square mesh biplane grids of circular cylinders at the entrance to the working section. In the absence of a mesh, the free stream turbulence (downstream component) is approximately 0·06 %.

The basic measuring instrument is the hot-wire anemometer, used in conjunction with frequency compensation circuits in the manner described frequently in the literature. Certain precautions are necessary in order that accurate measurements of the experimental quantities can be made, the most important being the use of wires as short as can be managed, in practice of lengths between 0·5 and 1 mm. The hot-wires used are of Wollaston wire 0·0001 in. diameter, and are operated at a temperature near 150° C. The compensation necessary to correct for heat lag is determined electrically using an interrupter.

* Full details of the design and characteristics of the wind tunnel are given in an unpublished paper by D. C. MacPhail, Royal Aircraft Estab., Report No. Aero. 1928, 1944.

All terms of equation (2 7) are completely determined by u', λ, S and G The measurement of u' is straightforward, but the remaining three quantities are determined by a method depending on the use of electrical differentiating circuits, and the assumption that

$$\frac{\partial u}{\partial x} = -\frac{1}{U}\frac{\partial u}{\partial \tau},$$

where x represents distance downstream from the mesh, τ is the time variable for the velocity fluctuation recorded by the hot-wire, and U is the mean stream velocity. This requires that the turbulent velocities be small compared with the mean stream velocity, so that the flow pattern may be considered as remaining unchanged while it is swept past the wire. From the expressions for λ, S and G in terms of spatial derivatives we have

$$\frac{1}{\lambda^2} = -f_0'' = \frac{1}{u'^2}\overline{\left(\frac{\partial u}{\partial x}\right)^2} = \frac{1}{u'^2 U^2}\overline{\left(\frac{\partial u}{\partial \tau}\right)^2},$$

$$S = -\lambda^3 k_0''' = -\overline{\left(\frac{\partial u}{\partial x}\right)^3}\Big/\left[\overline{\left(\frac{\partial u}{\partial x}\right)^2}\right]^{\frac{3}{2}} = \overline{\left(\frac{\partial u}{\partial \tau}\right)^3}\Big/\left[\overline{\left(\frac{\partial u}{\partial \tau}\right)^2}\right]^{\frac{3}{2}},$$

$$G = \lambda^4 f_0^{\mathrm{iv}} = u'^2\overline{\left(\frac{\partial^2 u}{\partial x^2}\right)^2}\Big/\left[\overline{\left(\frac{\partial u}{\partial x}\right)^2}\right]^2 = u'^2\overline{\left(\frac{\partial^2 u}{\partial \tau^2}\right)^2}\Big/\left[\overline{\left(\frac{\partial u}{\partial \tau}\right)^2}\right]^2,$$

and the practical application only depends on the possibility of performing the differentiations electrically, which is straightforward, and of measuring the mean cube and the mean square of a fluctuating voltage.

Electrical differentiation depends on the use of the charging current of a condenser to produce a potential drop in a series resistance proportional to the time derivative of the potential difference across the condenser. When substantially the whole applied potential difference appears across the condenser and very little across the resistance, then it is accurate to take the potential drop across the resistance as proportional to the time differential coefficient not only of the potential across the condenser but also of the applied potential The circuit used in these experiments (figure 2) is fundamentally similar, but uses two resistance-capacity stages and an amplifier stage to produce the convenient characteristics of unit amplification in the middle of the frequency range and an upper frequency limit of operation which is beyond the capability of a hot-wire anemometer For the measurement of the double derivative in G the above process is simply repeated on the output from the first differentiation.

It should be noted that this method of obtaining correlation derivatives from measurements of time derivatives of the velocity, although in principle capable of indefinite extension to higher order derivatives, cannot be so extended, because amplifier noise is a limiting factor, and indeed makes the measurement of the double derivative in G very difficult at low values of the intensity of turbulence. It would be impossible to measure higher order derivatives with the present arrangement, but if there were available an improved input arrangement having a lower noise level, e.g. an input transformer, then some extension might be possible.

Two methods of measuring the mean cube have been used, the first involving the determination of various moments of the statistical distribution of $\partial u/\partial r$ by numerical integration. This distribution is obtained from an electronic 'gate' circuit passing current only when the input voltage is within a prescribed narrow

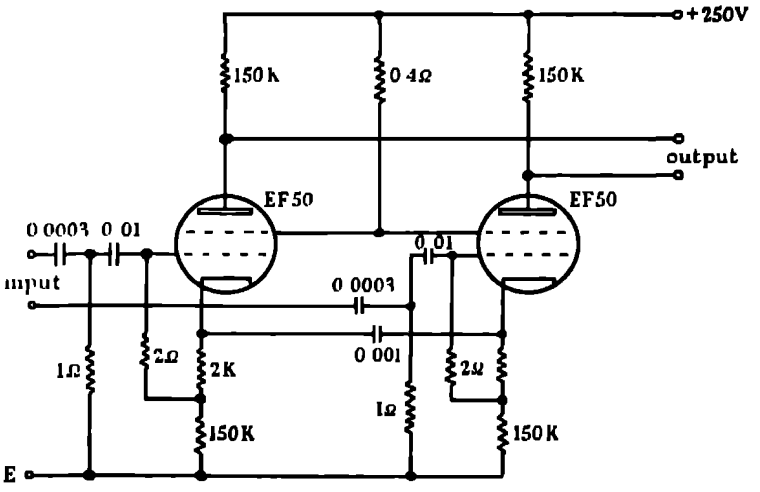


FIGURE 2. Differentiating circuit.

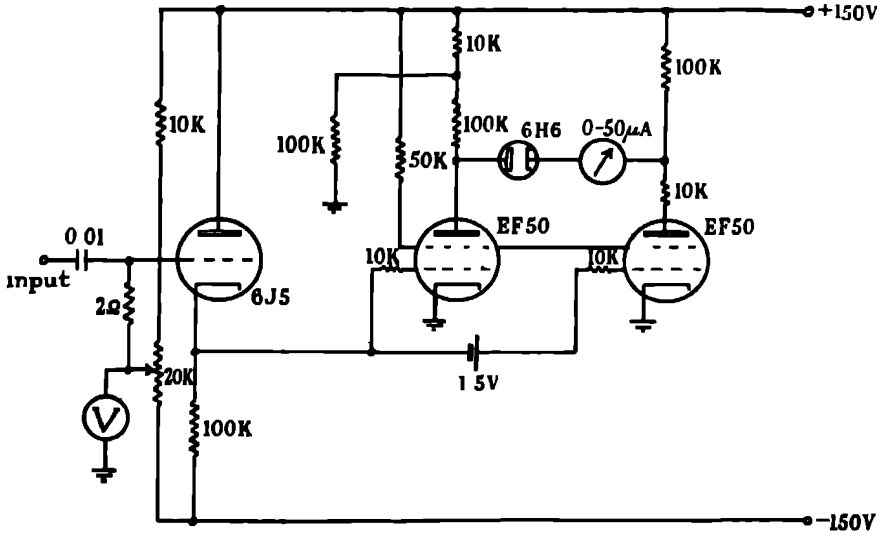


FIGURE 3. Statistical analyzer.

range By varying the position of this range on the voltage scale, and by measuring the mean current at each position by means of a long-period microammeter, the complete probability distribution of a fluctuating voltage may be determined.

The second method is purely electronic An output proportional to the instantaneous square of the input potential is obtained from a circuit depending on the curvature of the plate characteristic of a triode, and is multiplied by the original

input in a balanced modulator circuit, thus producing a current output proportional to the cube of the input voltage The mean value of this current is measured as before by a long-period microammeter The circuits of these two devices for measuring triple correlations are given in figures 3 and 4 and a more complete description will be published shortly. The first method is accurate but very tedious, while the direct measurement suffers from statistical fluctuations in the mean making precise observations difficult. The greater speed of reading makes the latter method preferable, and although similar results have been obtained by both methods, nearly all the results shown below are determined by the electronic method The measurement of mean squares is made with a thermo-junction and millivoltmeter.

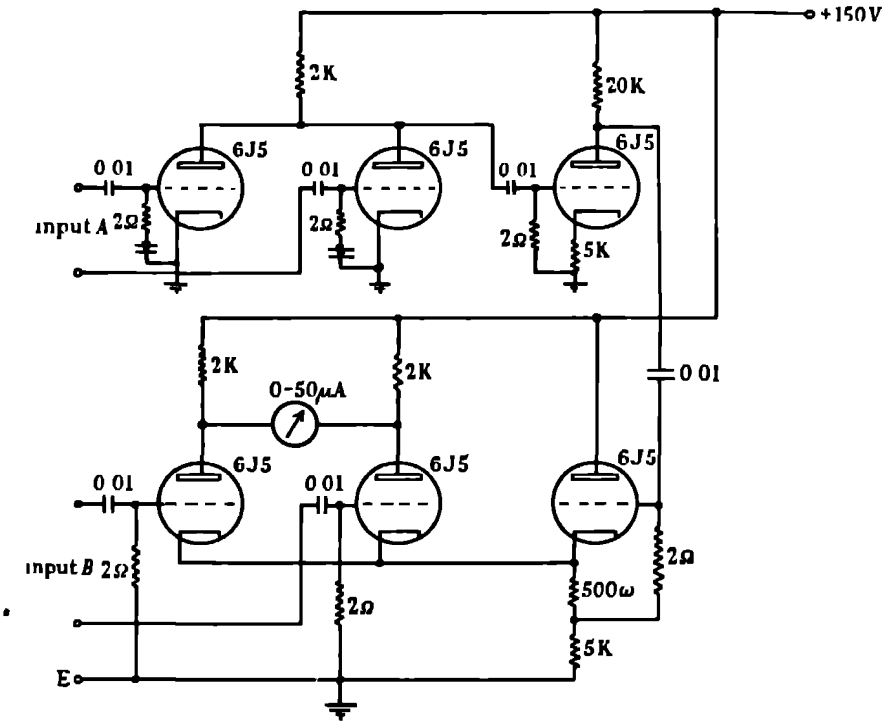


FIGURE 4. Triple correlator.

4 EXPERIMENTAL RESULTS

The measured values of u' are plotted in figure 5 in the form $(U/u')^2$ against x/M, where x is the distance of the point of measurement from the grid of mesh size M, and U is the velocity of the stream flowing past the grid As is usual, the time of decay in the idealized theoretical problem of homogeneous turbulence is identified with the quantity x/U occurring in the experimental turbulence which is slightly non-homogeneous The abscissa x/M is therefore the time decay made dimensionless by use of the factors M and U. Experiments were made over a range of speeds U = 640 cm /sec to 1280 cm./sec. and with three grids, M = 1·27, 2·54 and 5 08 cm.,

although only the two smaller mesh sizes were useful for the deduction of decay laws Figure 5 shows the experimental points to be approximately independent of M and U as noted by many previous workers, and to lie approximately on a straight line for the decay range $x/M = 20$ to 120 Further evidence that $(U/u')^2$ varies linearly with x/M will be obtained from the variation of λ^2 with x

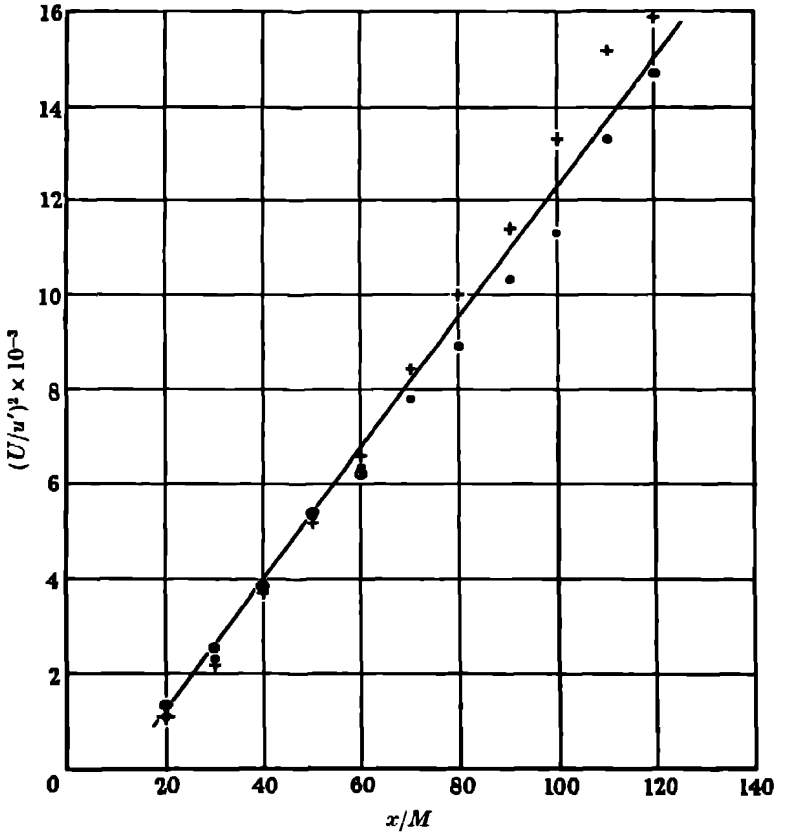


FIGURE 5 Decay of intensity of turbulence.

●, $U = 643$ cm. sec.$^{-1}$, $M = 1{\cdot}27$ cm , +, $U = 1286$ cm sec.$^{-1}$, $M = 1\ 27$ cm , ⊙, $U = 643$ cm. sec.$^{-1}$, $M = 2\ 54$ cm

Figure 6 shows the corresponding experimental values of λ^2 plotted against x. It is difficult to assess the accuracy of the measurements, but small values of λ^2 should be treated with caution owing to the effect of finite length of the heated wire. It seems reasonable to infer from the figure that λ^2 varies linearly with x over the decay range concerned, and that the variation does not depend on M.

Now the energy equation (1·5) can be written

$$\frac{d \log \frac{U^2}{u'^2}}{dt} = \frac{10\nu}{\lambda^2}. \tag{4·1}$$

Thus if $(U/u')^2$ is proportional to $(t-t_0)^n$ for some part of the decay range, the energy equation—which is an exact deduction from the condition of isotropy—demands for consistency that

$$\frac{d\lambda^2}{dt} = U\frac{d\lambda^2}{dx} = \frac{10\nu}{n}. \tag{4·2}$$

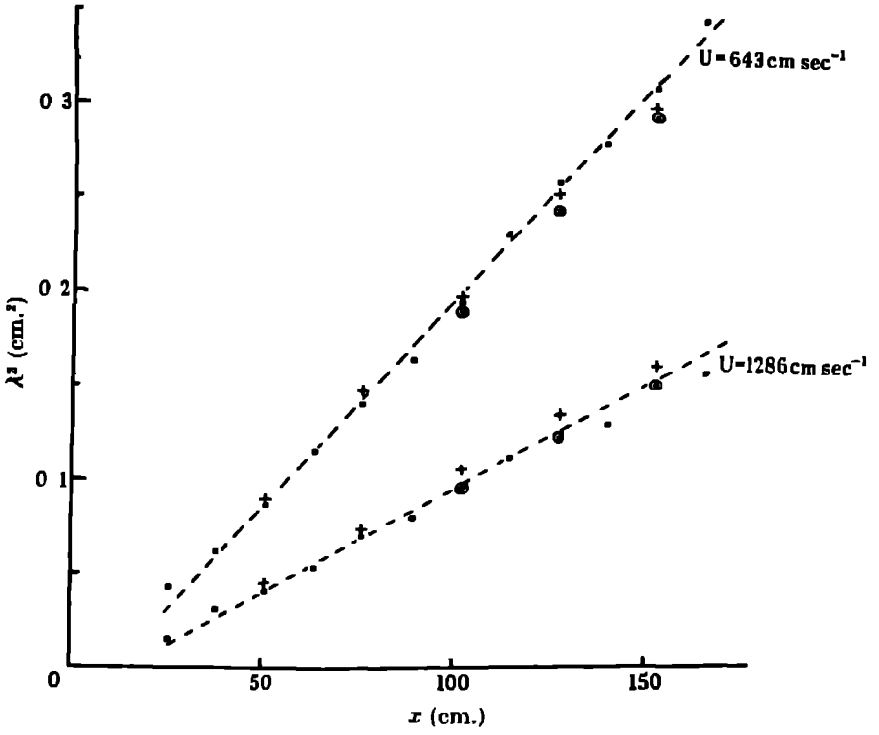


FIGURE 6 Variation of λ during decay

●, $M = 1{\cdot}27$ cm , +, $M = 2\ 54$ cm., ○, $M = 5\ 08$ cm ; ---, $\dfrac{d\lambda^2}{dx} = \dfrac{10\nu}{U}$.

Straight lines with the slope $10\nu/U$ are drawn through the two sets of experimental points in figure 6 and since the agreement is good, it is concluded that under the circumstances of the present experiments $n = 1$ and both $(U/u')^2$ and λ^2 increase linearly with the time of decay or distance from the mesh *

* Dryden (1943) has concluded from a survey of the available data that $(U/u')^2$ varies during decay as some power of t lying between one and two. He found the data from different sources not to be wholly consistent, and it is not possible to be certain about the reasons for these variations Previous measurements of U/u' have not been accompanied by measurements of λ^2 during decay, so that the check on the decay law provided by (4·2) has not hitherto been used. While the authors therefore accept the decay law $(U/u')^2 \propto t$ as applicable under the conditions of the above experiments ($UM/\nu = 5\ 5 \times 10^3$ to $4\ 4 \times 10^4$), it is without prejudice to the possibility of slightly different laws applying under other conditions. Indeed it seems impossible on theoretical grounds for the above decay law to hold for indefinitely large times of decay. The authors hope to take up the question of the decay of energy and of the form of the double-velocity correlation function in another communication.

It follows that the Reynolds number of turbulence given by R_λ, $= u'\lambda/\nu$, is approximately constant during decay (and accurately so if $(U/u')^2$ and λ^2 vanish at the same value of x) and the experimental values shown in figure 7 are in agreement with this approximate relation Using these results, it also follows that the vorticity equation (2·7) reduces to a simple form, thus

$$\frac{14}{3\sqrt{5}}\omega'^3\frac{G}{R_\lambda}-\frac{7}{3\sqrt{5}}\omega'^3 S=-\frac{d\left(\frac{5u'^2}{\lambda^2}\right)}{dt}$$

$$=-\frac{50\nu u'^2}{\lambda^4}+\frac{5u'^2}{\lambda^4}\frac{d\lambda^2}{dt}=\frac{100\nu u'^2}{\lambda^4},$$

i e.
$$G=\tfrac{10}{7}+\tfrac{1}{2}R_\lambda S. \tag{4·3}$$

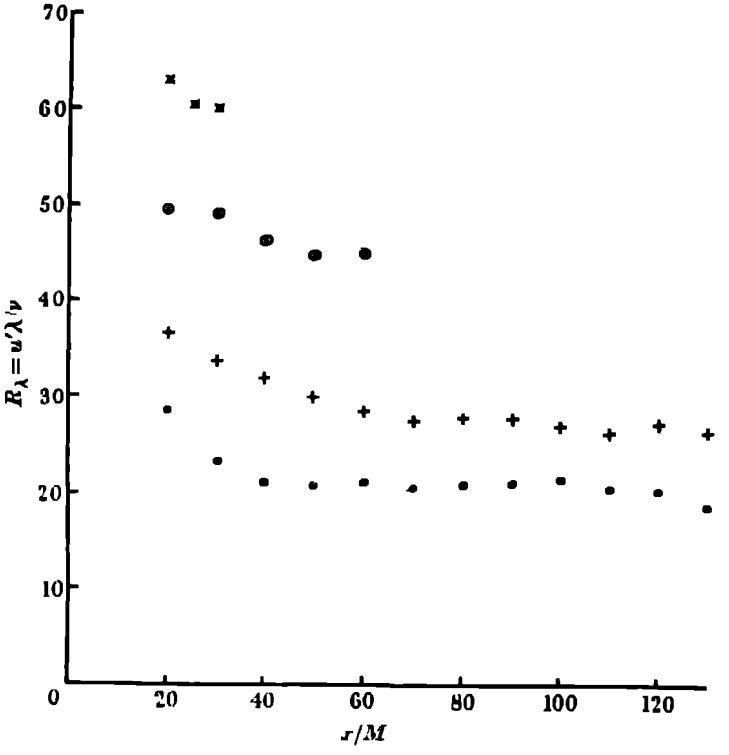


FIGURE 7. Variation of R_λ during decay.

●, $R_M = 5620$, +, $R_M = 11{,}250$, ⊙, $R_M = 22{,}500$, ×, $R_M = 45{,}000$.

Measurements of S at different distances from the grid are shown in figure 8 for particular values of M and U The inference from these, and other similar measurements, is that S does not change significantly over the decay range concerned In view of (2·4), the meaning of this result is that when r is small, the parabolic curve describing the function $f(r)$ and the cubic curve describing the function $k(r)$ maintain their shape and change their scale during decay by a similar factor

The average value of S over the range of x/M covered by the measurements in each case is plotted in figure 9 against the Reynolds number of the mesh R_M, $= UM/\nu$. Apart from changes during decay of the turbulence, and apart from experimental factors such as the background turbulence of the wind tunnel, this factor R_M uniquely determines S and all other dimensionless parameters of the turbulence;

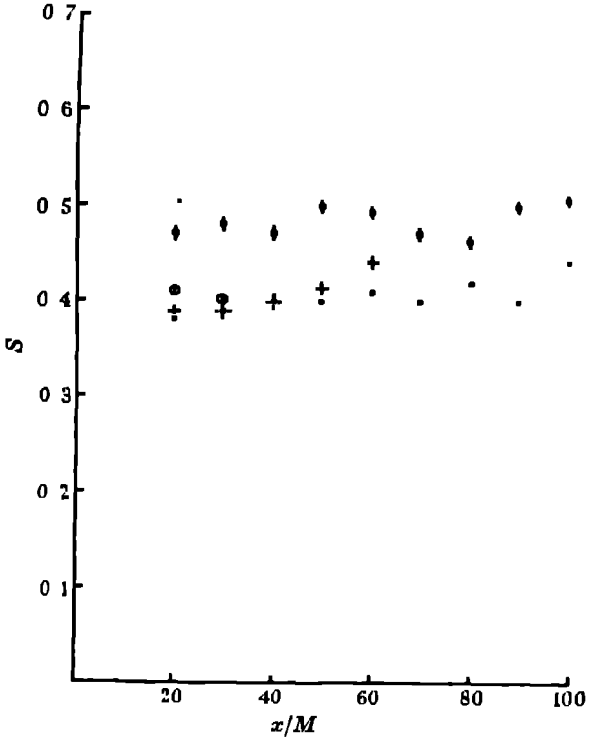


FIGURE 8 Variation of S during decay.

ϕ, $R_M = 5620$; •, $R_M = 7960$, +, $R_M = 15{,}820$, ○, $R_M = 31{,}640$.

figure 9 does in fact suggest that the points fall around a single curve. The experimental points appear to indicate a slight decrease in S as R_M increases, but in view of the large range of conditions represented in the figure and the consequent scope for systematic errors in the measurements, it would be unwise to be definite about this tendency. Failing more certain evidence to the contrary, it will be assumed that S is an absolute constant, independent of R_M over the range $5{\cdot}5 \times 10^3$ to $4{\cdot}4 \times 10^4$, with the value 0·39, which is the average of the S values of the points shown in figure 9.

Measurements of G at different stages of the decay process are shown in figure 10, and demonstrate that G may also be regarded as constant during decay. However, G is not independent of the mesh Reynolds number and in figure 11 the average value of G during the decay is plotted against R_λ, $= u'\lambda/\nu$. The Reynolds number R_λ is used here instead of R_M since the relation between G and R_λ is evidently

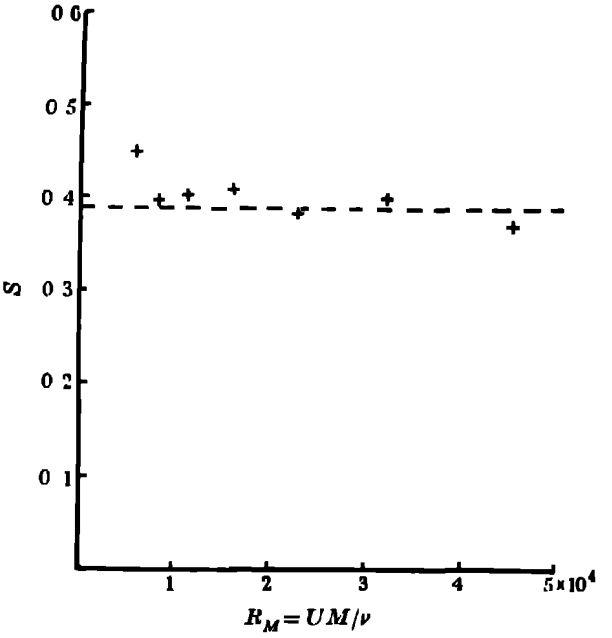


FIGURE 9. Variation of S with mesh Reynolds number.

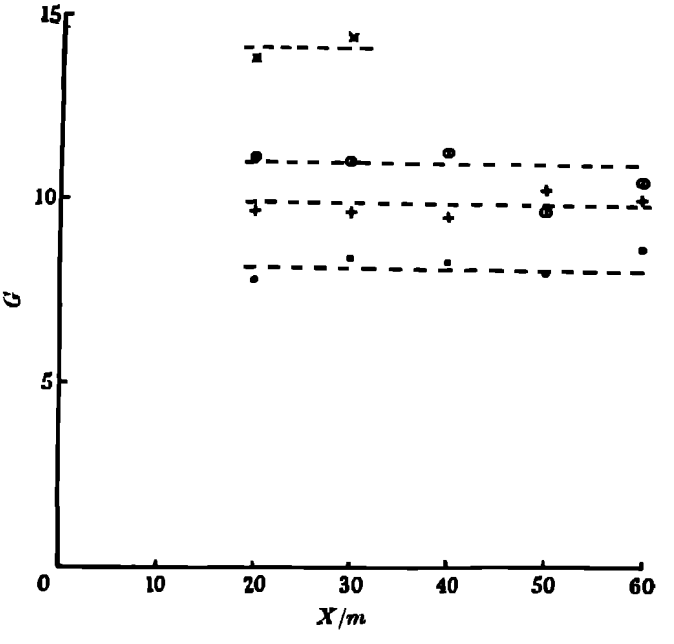


FIGURE 10. Variation of G during decay.

×, $M = 5{\cdot}08$ cm, $U = 909$ cm sec^{-1}; ○, $M = 2{\cdot}54$ cm, $U = 909$ cm sec.$^{-1}$;
+, $M = 2{\cdot}54$ cm, $U = 643$ cm sec.$^{-1}$; ●, $M = 1{\cdot}27$ cm, $U = 643$ cm.sec.$^{-1}$

approximately linear. The relation between the two Reynolds numbers is easily deduced from the experimental fact that

$$\frac{d\left(\frac{U}{u'}\right)^2}{d\frac{x}{M}}, \quad = 10\frac{R_M}{R_\lambda^2},$$

is approximately independent of R_M, i e. R_λ is proportional to $R_M^{\frac{1}{2}}$ The single point indicated in figure 11 is the value of G obtained by Taylor (1938) from measurements of the shape of the correlation function $f(r)$ near the origin, and is consistent with the present measurements

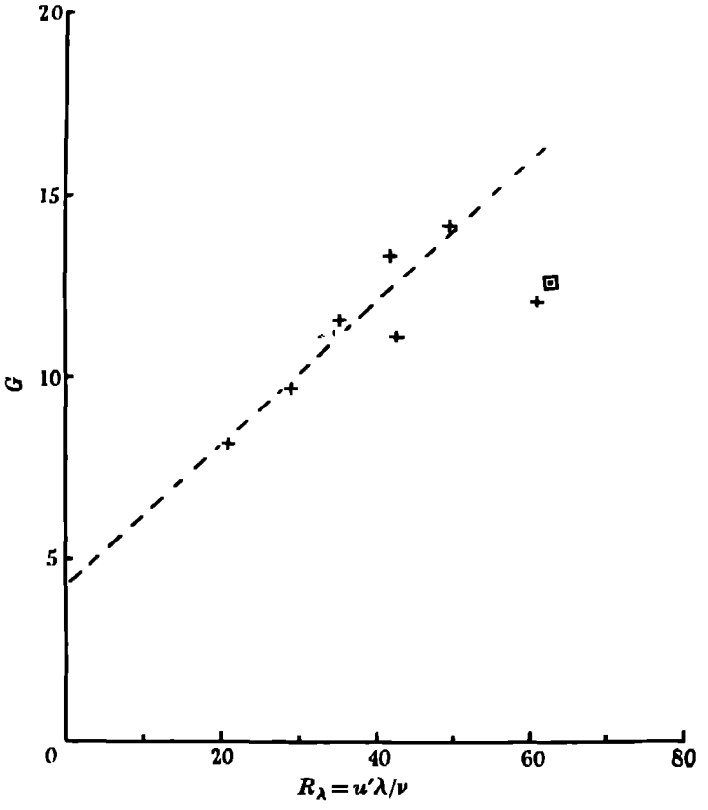


FIGURE 11. Variation of G with R_λ.
+, Experimental; □, Result of G. I Taylor, ---, $G = \frac{3}{7}0 + \frac{1}{2}R_\lambda S$ with $S = 0{\cdot}39$.

Also shown in figure 11 as a full curve is the value of G deduced from the measured value of S, viz. 0·39, and the vorticity equation (4·3) The agreement between this relation for G and the direct measurements is quite reasonable in view of the many separate measurements involved The exact value of S is of considerable interest, and to this end it is worth noting that a value of S other than 0 39 would not give an

agreement between the points and the line in figure 11 which is significantly better The general agreement with the vorticity equation (4 3) is a confirmation, in the first place, of the assumption of isotropy of the turbulence behind a grid and of the mathematical deductions from isotropy, and in the second place, of the consistency of the measurements of the four quantities u', λ, S and G.

5 Discussion of results

If the results described in the preceding section are accurate, there is available some very interesting information on which to base physical theories of the mechanism of turbulence It will always be wise, however, to keep in mind the conditions under which the information was obtained; the range of grid Reynolds numbers and the range of (non-dimensional) decay times are the important factors. Some of the more direct deductions from the experimental data are indicated here

(*a*) Since S, G and R_λ have all been shown to be approximately constant during decay, it follows from (2 7) that the rate of production of ω'^2 due to diffusive stretching of the vortex tubes, and the rate of dissipation of ω'^2 due to viscosity, are each proportional to ω'^3 during the decay process. The difference between these two rates gives $d\omega'^2/dt$ which is likewise proportional to ω'^3 and we must therefore have

$$\frac{d\left(\frac{1}{\omega'}\right)}{dt} = \text{constant}, \tag{5·1}$$

as indeed is required by the linear dependence of $(U/u')^2$ and λ^2 on t It also follows that the ratio of the rates of production and dissipation of ω'^2 is constant, and the vorticity balance is essentially the same at all stages of the decay This fact provides support for theories which maintain that there is a dynamical similarity, at different stages of the decay process, of those aspects of the turbulence which control the vorticity balance (Kolmogoroff 1941; Batchelor (in the press))

The ratio of the contribution to $d\omega'^2/dt$ from viscosity to that from vortex extension is $2G/R_\lambda S$, and using (4 3) with the value $S = 0\ 39$, this becomes

$$1+\frac{22}{R_\lambda}$$

The effect of viscosity is thus always the greater, and for the range of conditions of the above experiments, exceeds the effect of vortex extension by an amount ranging from 35 to 100 %

(*b*) In view of the attempts of v. Karman & Howarth (1938) and of Dryden (1943), to deduce the laws of decay of turbulence from the assumption that the various correlation functions $f(r)$, $k(r)$, etc , are functions only of r/l, where l is a length which may change during decay, it is of interest to note that the measurements show G, $= f_0^{iv}/(f_0'')^2$, to be constant during decay. Thus if $f(r)$ is expanded in powers of r^2, the terms as far as r^4 are of a form consistent with the assumption of preservation

of the shape of the double velocity correlation function during decay. From the experimental result $S = -k_0''' \lambda^3 =$ constant, it follows that the triple velocity correlation function also has the form $k(r/l)$, where l is the same length parameter as occurs in the double velocity correlation function.

(*c*) If the result that S is an absolute constant be assumed to hold for indefinitely large values of R_M (for theoretical support of this assumption, see Batchelor (in the press)), then at large Reynolds numbers equation (4·3) approximates to the form

$$G = \tfrac{1}{2} R_\lambda S = 0{\cdot}20 R_\lambda. \tag{5·2}$$

This approximate equation corresponds to the statement that the rates of dissipation and production of vorticity are both large compared with their difference. The point of interest is that for large values of R_M,

$$\lambda^4 f_0^{\mathrm{iv}}, \quad = \frac{f_0^{\mathrm{iv}}}{(f_0'')^2}, \quad \propto R_M^{\frac{1}{2}}. \tag{5·3}$$

This asymptotic dependence on R_M can be given a geometrical interpretation. For representing $f(r)$ near the origin by a limited power series in r^2,

$$f(r) = 1 - \frac{1}{2}\frac{r^2}{\lambda^2} + \tfrac{1}{24} f_0^{\mathrm{iv}} r^4, \tag{5·4}$$

and the slope at the point of inflexion of the curve given by (5·4) is

$$\left(\frac{\partial f}{\partial \dfrac{r}{M}}\right)_{f''=0} = -\frac{2}{3}\frac{M}{\lambda}\left(\frac{2}{\lambda^4 f_0^{\mathrm{iv}}}\right)^{\frac{1}{2}}. \tag{5·5}$$

Since M/λ is proportional to $R_M^{\frac{1}{2}}$, the expression on the right varies as $R_M^{\frac{1}{2}}$ when R_M is large The meaning of this result is that, as $R_M \to \infty$, the slope of the correlation curve immediately outside the region near the origin where it is parabolic increases indefinitely Since $\lambda/M \to 0$ as $R_M \to \infty$, the parabolic region becomes small and the correlation curve therefore becomes cusp-like at the origin as the Reynolds number becomes large

In a paper already quoted, Taylor (1938) made the contrary assumption that the slope of the correlation curve $f(r)$ immediately outside the parabolic region remains unchanged as the wind speed U increases From this it was deduced that $\lambda^4 f_0^{\mathrm{iv}}$ is proportional to U and that $\nu u'^2 f_0^{\mathrm{iv}}$ is proportional to U^5. Since the total rate of change of ω'^2 can be written

$$\frac{d\omega'^2}{dt} = \frac{U^4}{M^2}\left(d\left(5\frac{u'^2}{U^2}\frac{M}{\lambda^2 U}\right)\Big/ d\frac{x}{M}\right)$$
$$\propto M^{-2} U^4,$$

Taylor concluded that the ratio of the rate of dissipation of vorticity to total rate of change in ω'^2 varies as U when U is large The result of the present paper is (from

(5·3)) that $\nu u'^2 f_0^{\frac{1}{2}\nu}$ is proportional to $M^{-1\cdot5}U^{4\cdot5}$, and thence that the two contributions to $d\omega'^2/dt$ are each of order $R_M^{\frac{1}{2}}$ higher than their resultant when R_M is large, as in fact is evident from (4·3)

It is a pleasure for the authors to acknowledge the stimulation provided by frequent conversations with Professor G. I. Taylor during the course of the research. The work was carried out while one of us (G. K. B.) was the holder of an Aitchison Travelling Scholarship from the University of Melbourne, and the other (A. A. T.) was the holder of an Overseas Scholarship of the Royal Commission for the Exhibition of 1851 and a scholarship provided by Emmanuel College. Both authors received supplementary grants from the Science and Industry Endowment Fund, Australia. Component parts of the apparatus for measuring the turbulence were purchased with money from a grant provided by the Aeronautical Research Council.

References

Batchelor, G. K. *Proc. Camb. Phil. Soc.* (in the Press)
Dryden, H. L. 1943 *Quart. Appl. Math.* **1**, 7–42.
Karman, T. v. 1937 *J. Aero. Sci.* **4**, 131–138.
Karman, T. v. & Howarth, L. 1938 *Proc. Roy. Soc.* A, **164**, 192–215.
Kolmogoroff, A. N. 1941 *C. R. Acad. Sci. U.R.S.S.* **30**, 301–305.
Taylor, G. I. 1935 *Proc. Roy. Soc.* A, **151**, 421–478.
Taylor, G. I. 1938 *Proc. Roy. Soc.* A, **164**, 15–23.

Measurements in the turbulent wake of a cylinder

By A A Townsend, *Emmanuel College, Cambridge*

(*Communicated by Sir Geoffrey Taylor, F.R.S.—Received 17 December 1946*)

An experimental study has been made of the conditions in the turbulent wake behind a cylinder in an air stream at cylinder Reynolds numbers between 100 and 7000. The measurements include the mean velocity distribution, the distribution of the three components of turbulence, the statistical distribution in time of the turbulent velocity components, and the correlation derivatives in the downstream direction for both the longitudinal and transverse correlations. It is shown that the intensities of the three components of turbulent velocities are nearly equal except at the centre of the wake, where the balance between viscous decay and diffusion from the regions of shear results in a considerably greater intensity for the component at right angles to the direction of flow and also to the axis of the cylinder, i.e. in the direction of diffusion of turbulent energy. An analysis of the energy in the wake proves that complete dynamical similarity in the wake, if it is ever attained, does not occur closer to the cylinder than 1000 diameters, and that the viscous dissipation of energy can be described in terms of a length parameter, nearly constant over the width of the wake and approximately equal to the measured value of Taylor's scale of micro turbulence.

Introduction

Theoretical descriptions of conditions in non-uniform turbulent flow have been made, usually in terms of mixture-length theories which are based on the analogy between the turbulent motion of a fluid and the random motion of gas molecules. It is evident that while the assumptions made in these theories may lead to predicted mean-velocity distributions agreeing closely with observed distributions (Goldstein 1938), because of the discontinuous nature of the assumed processes, it is impossible to obtain from mixture-length considerations any description of the turbulent motion in the wake, apart from the eddy shearing stress and the diffusive properties. This emphasis on the properties of the mean flow is a reflexion of the comparative lack of information on the turbulent intensities in the wake, and the only theoretical investigation using the methods of the statistical theory of turbulence is that of Hu (1944). To solve the problem, Hu uses many assumptions which though mathematically convenient are not necessarily accurate, and more detailed information on the turbulence is necessary to guide such investigations. The object of this paper is to record a fairly detailed study of the fully developed turbulent wake behind a circular cylinder in a low-turbulence air stream. Measurements of the intensities of the three components of turbulence have been made over a considerable part of the wake, and some measurements of correlations and statistical distributions of velocity components are also included.

Experimental method and equipment

The measurements were made in a small wind tunnel of the closed return type, using a fine gauze followed by a settling length and a rapid contraction to obtain a low-turbulence air stream. At the entrance the working section is 15×15 in., and

extends for 84 in. The average value of the downstream component of turbulence is 0·07 % and of the cross-stream component 0·14 %. Further details of the tunnel are given by MacPhail (1946).

Turbulent velocities were measured in the usual way with a hot-wire anemometer in conjunction with an amplifier compensating for the time lag in response of the hot-wire. For measurements of the downstream component, a simple wire at right angles to the air stream was used, 0·00025 cm diameter and about 1 mm. long. For cross-stream components, pairs of wires of similar dimensions were used, arranged in the form of a V. Such wires are easier to construct than the X-form, but possibly are slightly less accurate, due to the larger sensitive volume, and consequent errors in the measurement of small-scale turbulence. Intensity measurements are made in the usual way with a thermo-junction and a millivoltmeter.

Correlation derivatives are measured from the analysis of a hot-wire output by an electrical differentiating circuit. Since the turbulent velocities are small compared with the stream velocity, one may assume

$$\frac{\partial u}{\partial x} = -\frac{1}{U}\frac{\partial u}{\partial t},$$

where u is a velocity component at a fixed point, x is measured parallel to the mean stream velocity, U is the mean stream velocity, t is time. It is easily shown that if $R_u(x)$ is the correlation function between the instantaneous values of u at positions x apart on a line parallel to the direction of the mean stream,

$$\frac{\partial^2 R_u}{\partial x^2} = -\frac{\overline{\left(\frac{\partial u}{\partial x}\right)^2}}{\overline{u^2}}$$

and so

$$\frac{\partial^2 R_u}{\partial x^2} = \frac{1}{U^2}\frac{\overline{\left(\frac{\partial u}{\partial t}\right)^2}}{\overline{u^2}}.$$

Electrical differentiation with respect to time of a fluctuating potential depends on the relation that the charging current of a condenser is proportional to the time derivative of the potential difference across the condenser, and the production of a voltage proportional to this current across a resistance in series with the condenser (Townsend 1947).

To measure the statistical distribution in time of a velocity component, an electronic analyzer has been used. A valve circuit is arranged so that current is passed through a microammeter only when the applied voltage is within narrow fixed limits. If this applied voltage is the sum of a variable but known steady potential and the appropriate anemometer output, then the mean microammeter current is a measure of the probability that this sum should lie within the limits, and the complete distribution may be obtained by varying the steady potential (Townsend 1947).

Results

Intensity measurements of turbulence were made over a range of cylinder Reynolds numbers from 100 to 7000, but, as within this range the turbulence distributions were essentially similar sufficiently far downstream, comprehensive measurements have only been made in the wake of a 1·12 mm. diameter cylinder, at airspeeds of 560 and 1120 cm.sec $^{-1}$ Specimen distributions of the root-mean-square intensities of the three components and of the mean velocity across the wake

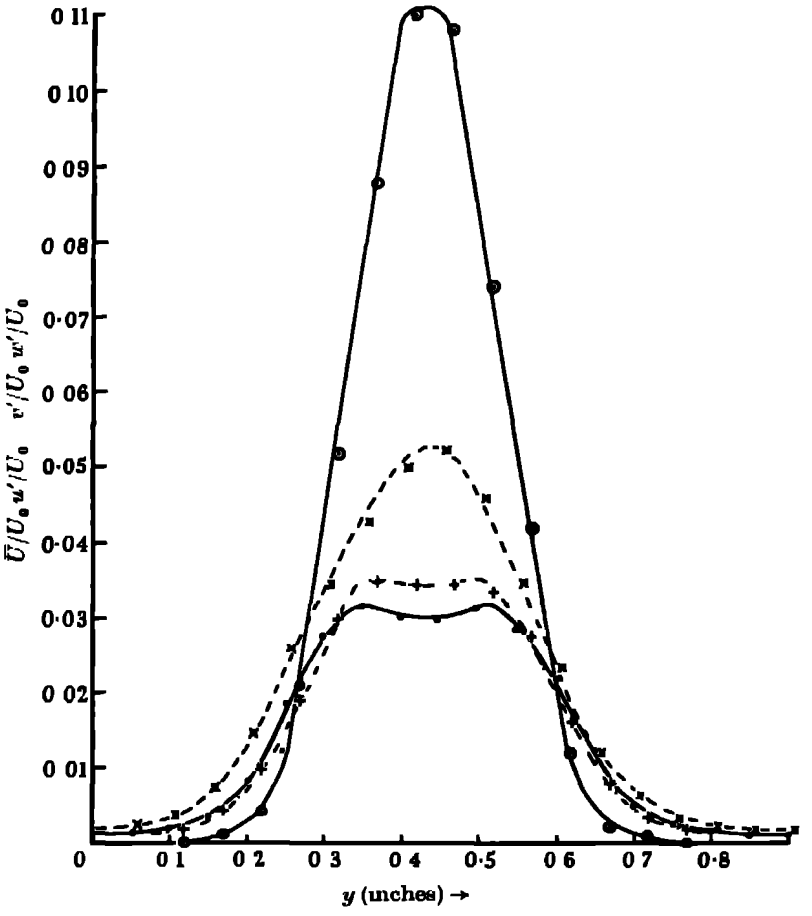


Figure 1. Distribution at 91 diameters.

are plotted in figures 1 to 3 The direction of u is taken as parallel to the mean flow, of v at right angles to the plane of the wake, and of w parallel to the axis of the cylinder The rapid development and stabilization of a characteristic shape of distribution is evident.

The statistical distributions in time of the velocity fluctuations have been determined, and it is found that except close to the edge of the wake, the distributions

are very nearly Gaussian as in isotropic turbulence The distribution of u at the edge of the wake is skew, as is easily seen in figure 4 In this diagram, for each curve the vertical and horizontal scales have been arbitrarily adjusted for easy comparison of the forms.

To establish a length connected with the scale of turbulence, measurements of $\partial^2 R_u/\partial x^2$ and $\partial^2 R_v/\partial x^2$ have been made in the wake In figure 5, the variation across the wake is represented, and both these quantities are seen to be nearly constant

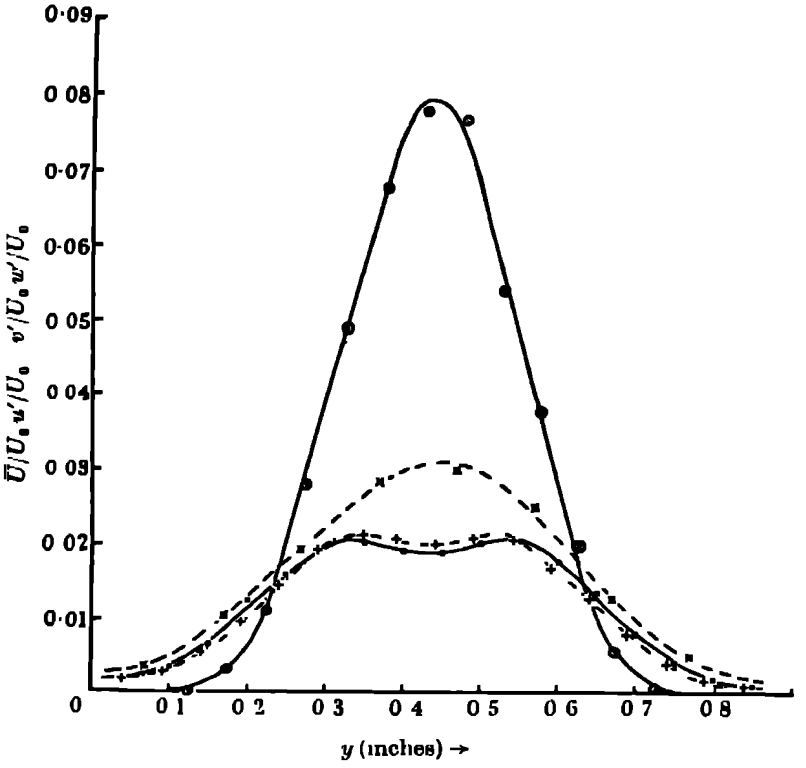


FIGURE 2. Distribution at 180 diameters.

over the whole wake, $\partial^2 R_v/\partial x^2$ being nearly twice $\partial^2 R_u/\partial x^2$, as it would be if the turbulence were isotropic. In figure 6, the variation of three lengths associated with the wake at a fixed distance downstream are plotted against distance, that is, λ_u, d and l, where $1/\lambda_u^2 = -\partial^2 R_u/\partial x^2$, d is the half-width of the u' distribution, l is twice the half-width of the mean velocity distribution.

For a comparison with the theory of complete similarity, the squares of these lengths are plotted against x, the distance downstream from the cylinder

Briefly, the results may be summarized

(a) Beyond 100 diameters downstream, a characteristic distribution of turbulence across the wake is attained, marked by a pronounced central minimum in

the u' and w' distributions but not in the v' distribution, and approximate equality of these components outside the central region

(*b*) The 'microscale of turbulence' represented by λ_u is approximately constant over the greater part of the width of the wake

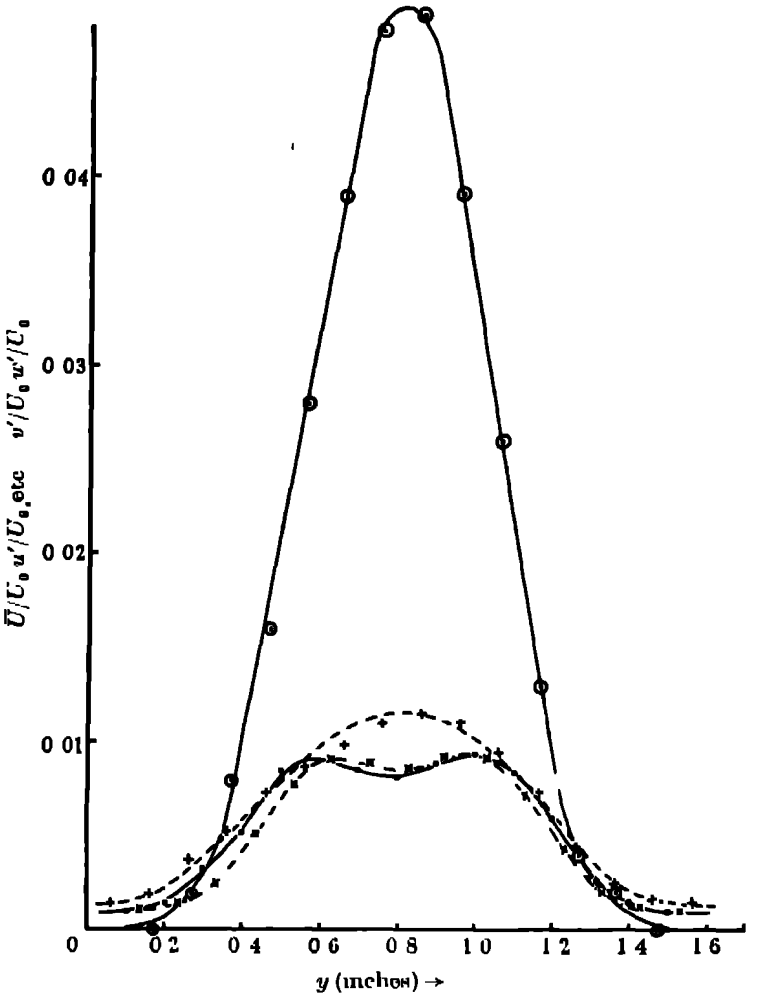


FIGURE 3 Distribution at 725 diameters.

FIGURES 1 to 3 $U_0 = 1120$ cm sec $^{-1}$, $D = 1.12$ mm.
⊙, mean velocity; ●, u component, +, v component ×, w component

(*c*) The statistical distributions in time of the velocity components are nearly Gaussian except that of u at the extreme edge of the wake

(*d*) The turbulence Reynolds's number based on the mean turbulent velocity and the width of the wake is approximately one-third of the cylinder Reynolds number.

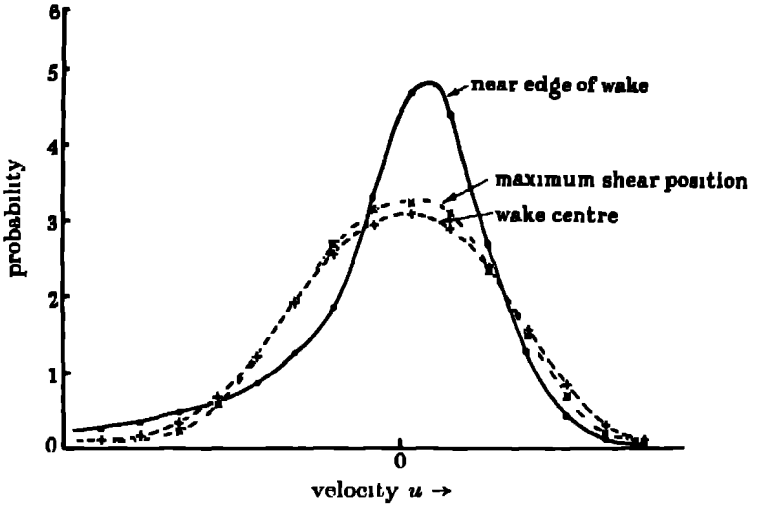


FIGURE 4 Statistical distribution of turbulent velocity

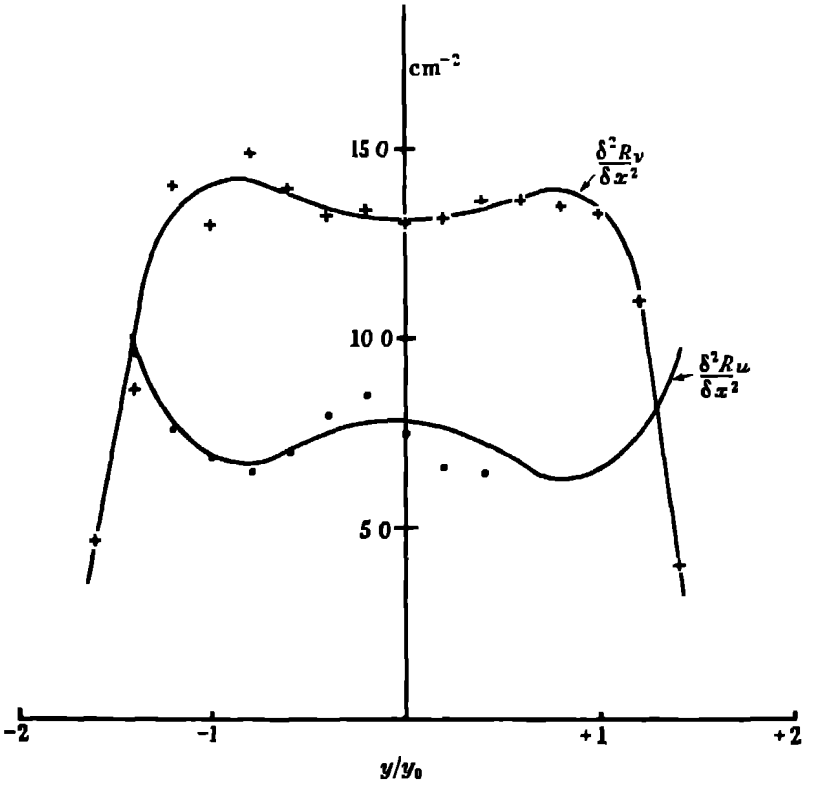


FIGURE 5. Variation of correlation derivatives in the wake.

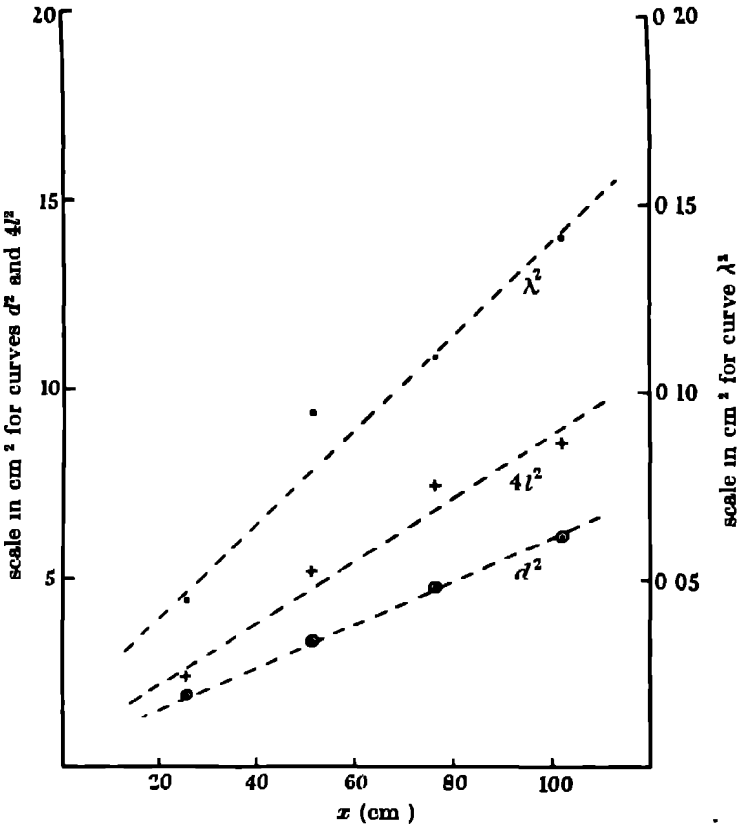


FIGURE 6. Change in breadth of the wake.

DISCUSSION OF RESULTS

From these results it is possible to calculate the total kinetic energy of the turbulent motion and the total kinetic energy of the mean flow in the wake at various points In figure 7 the variation of the integrals, I_T and I_M, with position is represented, where

$$I_T = \int_{-\infty}^{\infty} \frac{\overline{u^2}+\overline{v^2}+\overline{w^2}}{U_0^2}\,dy \quad \text{(curve I)},$$

$$I_M = \int_{-\infty}^{\infty} \frac{(U_0-\bar{U})^2}{U_0^2}\,dy \quad \text{(curve II)}$$

are measures of the energy in the mean stream and in the turbulent motion. Initially, I_T is larger than I_M, but the relative proportions of turbulent and mean flow energy change continuously, approaching in the limit the ratio $I_M/I_T = 4{\cdot}5$ (figure 8). This final value is only attained very slowly Complete equilibrium only occurs at distances in excess of 1000 diameters downstream. In view of this, it is rather surprising that properties of the mean flow appear to be stabilized after possibly 100 diameters, where the present experiments show that conditions are

rapidly changing (Goldstein 1938) It is evident that although the wake is not similar in detail in the region 100 to 1000 diameters downstream, some mean property must exist to replace the postulated turbulence field used in the transport theories A suitable quantity, as far as the numerical variation is concerned, is the total kinetic energy $(I_T + I_M)$, which is proportional to $x^{-\frac{1}{2}}$, and λ_u behaves in a suitable way to replace the arbitrary mixing length, i e it is proportional to $x^{\frac{1}{2}}$. It is not suggested that the correspondence of $(I_T + I_M)$ with the assumed turbulent intensity is exact, but the correspondence over the transition period is rather striking

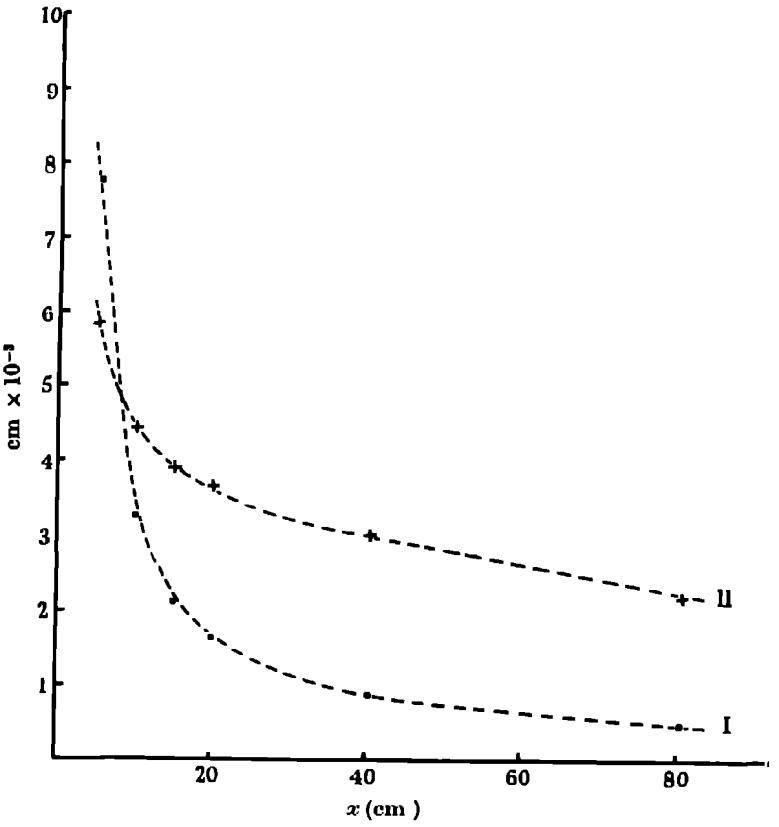


FIGURE 7. Dissipation of energy in the wake Wire diameter 1 12 mm , $U_0 = 560$ cm sec $^{-1}$.

By analogy with the equation for the dissipation of energy in isotropic turbulence, one may write

$$\frac{d}{dt}(I_T + I_M) = -10\nu \frac{I_T}{\lambda_e^2},$$

where λ_e is an effective 'scale of microturbulence'. Then λ_e may be calculated from the relation

$$\lambda_e^2 = 20\nu \left(\frac{I_T}{I_T + I_M} \right) \frac{x}{U}$$

deduced, using the experimental knowledge that the magnitude of $(I_T + I_M)$ is proportional to $x^{-\frac{1}{2}}$ λ_ϵ is then found to be approximately equal to λ_u, as shown in figure 10

From these measurements, it is possible to calculate the shear stress across the wake, and then to calculate correlation coefficients, R_{uv}, across the wake The results for one value of x are presented in figure 9

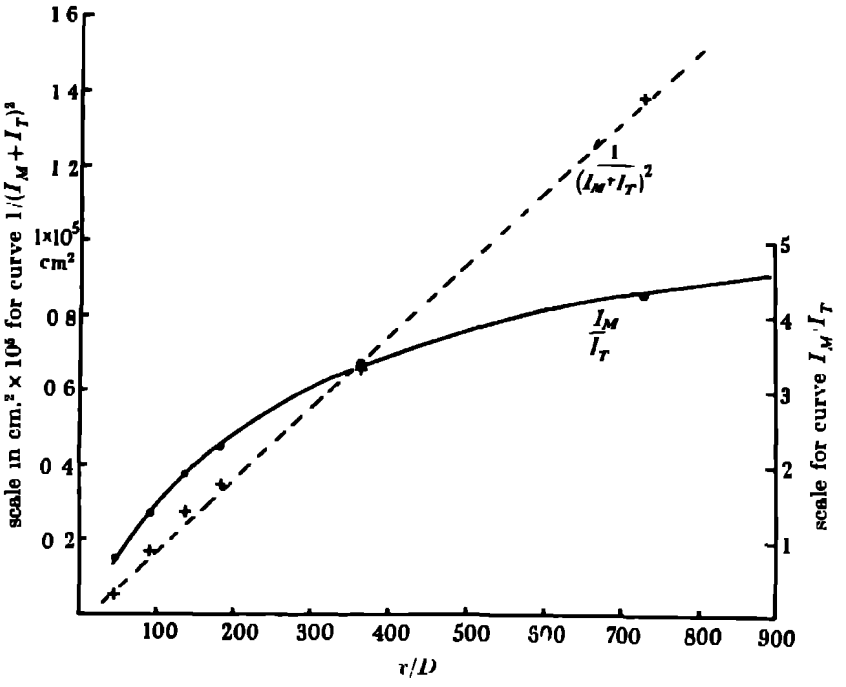


FIGURE 8 Energy relations in the wake $D = 1\ 12$ mm , $U = 1120$ cm sec $^{-1}$

The particular distribution of the velocity components of turbulence across the wake found in the experiments may be made to appear reasonable on qualitative grounds The processes occurring are the conversion of the energy of mean flow into turbulent energy by diffusion along the velocity gradients on each side of the wake centre, and the diffusion of the turbulent motion inwards toward the centre of the wake and outwards toward the undisturbed fluid, these parts of the fluid forming sinks of turbulent energy Thus the turbulent intensity at the wake centre is derived solely from diffusion of turbulent energy from the regions of shearing motion This diffusion takes place by movements of small bodies of fluid in the y direction, and since transport of turbulent momentum itself demands the presence of momentum in the y direction, but not necessarily in any other direction, it is not unlikely that the turbulent momentum diffusing from the regions of energy production should have a predominant component in the direction of diffusion The transfer of momentum components at right angles to the direction of diffusion will depend on the degree

of association of these components with the diffusive momentum component, and the tendency toward isotropy of the turbulent field. Now in the wake, diffusion toward the centre is carried out by movements in the y direction, and consequently the transport of v fluctuations will occur more rapidly than the transport of u or w

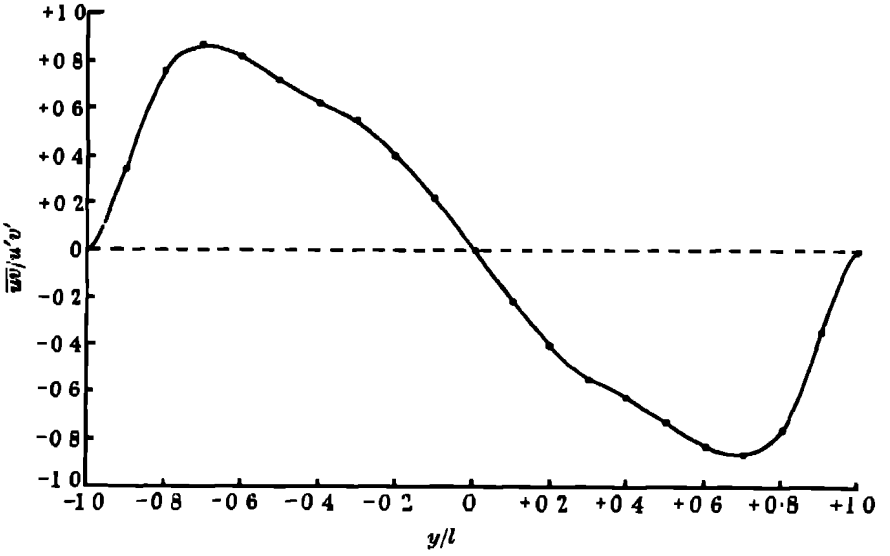


FIGURE 9 Variation of correlation coefficient

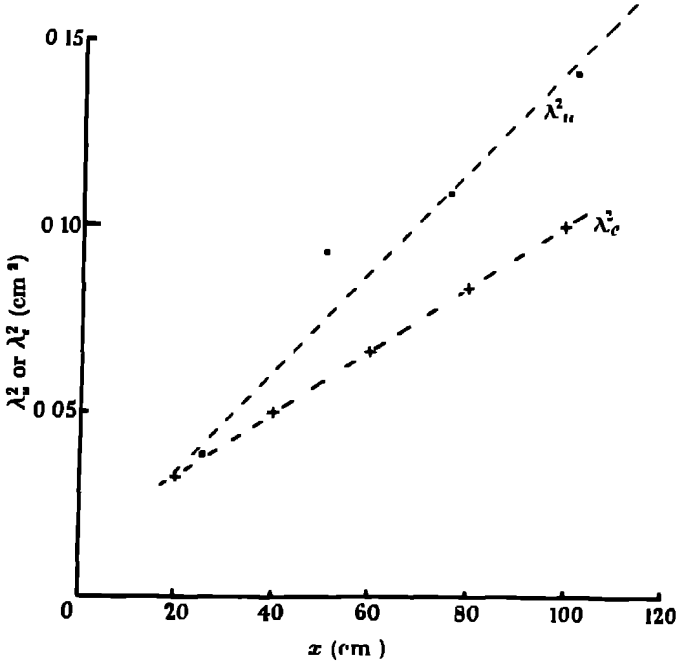


FIGURE 10. Variation of λ_u and λ_v. $U_0 = 1120$ cm sec^{-1}, $D = 1\ 12$ mm.

fluctuations, and it is therefore not unreasonable that at the wake centre the intensity of v fluctuations should be considerably larger than the intensities of u and w fluctuations Similar arguments show that the breadth of the v' distribution must be greater than that of either the u' or the w' distribution, a result also observed in these experiments.

Conclusions

The turbulence in the wake of a cylinder is rather more complex than has usually been assumed in existing theory, and the transition distance during which the final stable configuration is approached extends for nearly 1000 diameters Partial stability occurs after 100 diameters for the characteristics of the mean flow, but this is not reflected in the turbulent field, which remains in the transition state In the final state, the turbulent motion is small compared with the mean flow

The theoretical investigation of Hu, although it predicts velocity and temperature distributions with reasonable accuracy, does not explain the observed distribution of turbulent intensities across the wake.

An interesting coincidence is the approximate equality of λ_v and λ_w, which is particularly significant in view of the notion of local isotropy used by Kolmogoroff (1941 *a*, *b*), in his recent work on correlations in turbulent flow This would imply that the assumptions made by Kolmogoroff may be valid for the wake

It is evident that the conditions in the wake are more complicated than has been assumed, and that further experimental study is necessary for a complete understanding of the processes involved.

I wish to thank Sir Geoffrey Taylor for his advice and encouragement during the course of this investigation I am also indebted to the Royal Commission for the Exhibition of 1851 for an Overseas Scholarship, and to the Aeronautical Research Committee for a grant towards the cost of the turbulence measuring equipment

References

Goldstein, S 1938 *Modern developments in fluid mechanics* Oxford Clarendon Press
Hu, N. 1944 *Chinese J. Phys.* **5**, 30
Kolmogoroff 1941 *a* *C R Acad. Sci U R S.S* **30**, 301.
Kolmogoroff 1941 *b* *C.R. Acad Sci U R S S* **32**, 16
MacPhail D. C , 1944 *Royal Aircraft Estab Aero Rep.* No 1928
Townsend, A A. 1947 In course of publication.

The vibration spectrum and molecular configuration of 1 : 4-dioxane

By D A. Ramsay, *Laboratory of Colloid Science, University of Cambridge*

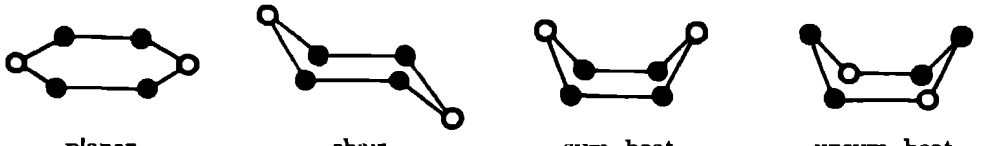
(*Communicated by E K Rideal, F.R.S —Received* 1 *January* 1947)

The symmetry properties and selection rules for the normal vibrations of 1 4-dioxane have been derived for all possible symmetries of the molecule The numbers of frequencies observed in the infra-red spectrum of the vapour and Raman spectrum of the liquid are shown to be inconsistent with structures of low symmetry While the number of observed Raman frequencies do not distinguish between the 'chair' and planar configurations, the number of observed infra red frequencies slightly favours the 'chair' structure The twelve vibration frequencies of the ring have been calculated for the 'chair' structure using a simple valency force field with four force constants Seven of the calculated frequencies agree with the observed frequencies to within 1 %, and two frequencies agree to within 3 to 4 %. Of the three remaining frequencies, two are active in the infra-red spectrum but are too low to be observed with prism spectrometers, while the third is active in the Raman spectrum but has not been experimentally observed. These three frequencies, however, agree satisfactorily with the corresponding assignments for cyclohexane The stretching constants of the C-C and C-O bonds, viz 3.7×10^5 and 4.45×10^5 dynes/cm , were found to be the same as for cyclohexane and dimethyl ether respectively

Introduction

In an earlier communication (Ramsay & Sutherland 1947) the physical evidence for the structure of the cyclohexane molecule was shown to favour a 'chair' configuration, and a satisfactory interpretation of the ring frequencies of the molecule was obtained by applying a simple valency force field In this paper the results of a similar investigation of the structure of the dioxane molecule are presented

The five possible ring structures for the dioxane molecule are (i) a planar configuration with symmetry D_{2h}, (ii) a 'chair' or *trans*-configuration with symmetry C_{2h}, (iii) a symmetrical 'boat' or *cis*-configuration with symmetry C_{2v}, (iv) an unsymmetrical 'boat' or *cis*-configuration with symmetry C_2, (v) an irregular configuration with no symmetry elements



○, oxygen atom ●, carbon atom

Figure 1

X-ray data for the structure of the dioxane molecule appear to be lacking, but the electron diffraction data of Sutton & Brockway (1935) showed that the molecule exists predominantly in the 'chair' form. Assuming all the carbon-bond angles to be tetrahedral and the C-C bond length to be 1·54 A, they obtained values of 1.46 ± 0.04 A for the C-O bond length and $110 \pm 5°$ for the oxygen-bond angle.

The dipole moment of dioxane has been measured by several workers. The early measurements of Sängewald & Weissberger (1929), Williams (1930), Smyth & Walls (1931), Hunter & Partington (1933) and Böeseken, Tellegen & Henriquez (1935), gave values ranging from 0 to 0·45 D which, however, were indistinguishable from zero by the methods used. Schwingel & Greene (1934) showed that the total molecular polarization of dioxane vapour was independent of temperature, indicating a zero dipole moment for the molecule, but these results were not confirmed by Kubo (1936) Later measurements by Vaughan (1939), using the liquid, however, confirmed the conclusions of Schwingel & Greene (1934) Hence the dioxane molecule must exist predominantly in the highly symmetrical planar or 'chair' configurations, and the less symmetrical structures, if present, must exist in very small proportions

The different symmetries of the above structures enable them to be distinguished further by the different numbers of fundamental vibration frequencies active in the infra-red and Raman spectra and by differences in the states of polarization of the Raman lines. This spectroscopic evidence for the configuration of the dioxane molecule is discussed in the following sections An attempt is then made to calculate the vibration frequencies of the dioxane ring using a simple valency force field and hence derive information concerning the strengths of the C-C and C-O bonds in the molecule

Experimental

The infra-red absorption of dioxane between 1500 and 700 cm $^{-1}$ in the vapour and in the liquid was measured using a Hilger D 209 double-beam spectrometer with a rock-salt prism. The dioxane was dried over sodium and distilled through a Fenske column, the constant-boiling middle fraction being used The vapour absorption was measured using a 25 cm. cell with rock-salt windows Liquid dioxane was contained in a small side bulb, the vapour pressure in the cell being determined by the temperature of the bulb. To prevent condensation of dioxane vapour on the rock-salt windows, the cell was heated electrically and maintained at a higher temperature than the bulb The absorption curves for different pressures of dioxane vapour are given in figure 2 and the frequencies are listed in table 1. Pressures up to 200 mm Hg were investigated, but the only additional band which began to appear was a very weak 740 cm $^{-1}$ frequency

The absorption of liquid dioxane was measured using a rock-salt cell of thickness 0·01 mm. The frequencies agreed well with those observed in the vapour but were slightly displaced to lower frequencies By comparing the absorptions of the liquid and the vapour it is seen that the 1042 and 1066 cm $^{-1}$ frequencies in the spectrum of the latter are P and R branches of a band with a strong central Q branch at 1054 cm $^{-1}$ Similarly, the vapour bands at 1450 and 1461 cm $^{-1}$ are due to rotational contour of a single band. The 880 and 889 cm.$^{-1}$ frequencies in the vapour spectrum, however, are distinct vibrational frequencies corresponding to the two frequencies at 874 and 887 cm $^{-1}$ in the liquid spectrum By increasing the absorption path of the liquid to 0·25 mm four additional weak frequencies were observed, viz 1020,

753, 735 and 711 cm $^{-1}$. The 735 cm.$^{-1}$ band corresponds to the 740 cm.$^{-1}$ frequency observed in the vapour spectrum at higher pressures That these bands were not due to small traces of impurities was confirmed by using an independent sample of dioxane when the frequencies again appeared with equal intensity.

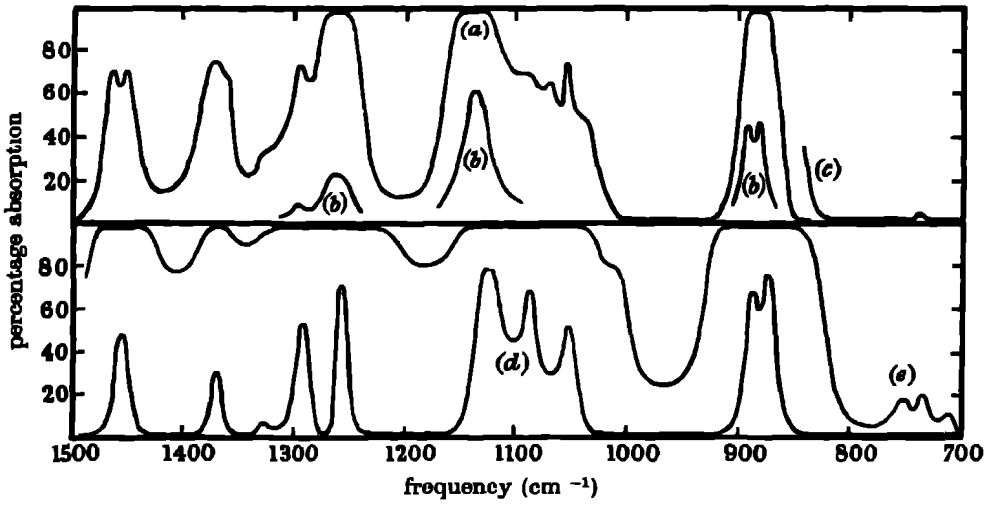


FIGURE 2. Vapour (*a*) pressure 30 mm. Hg, temp. 30° C, (*b*) pressure ~ 3 mm Hg, temp 20° C; (*c*) pressure 200 mm. Hg, temp. 100° C. Liquid (*d*) 0·01 mm. cell; (*e*) 0·25 mm. cell.

The frequencies listed in table 1 agree to ± 3 cm $^{-1}$ with the earlier measurements of McKinney, Leberknight & Warner (1937). These authors, however, did not observe the weak bands between 700 and 800 cm $^{-1}$ and did not record the 1327 cm $^{-1}$ vapour band as a separate frequency though their curves show a shoulder in this position They investigated the region 2·27 to 3·72μ with a prism-grating spectrometer of high resolving power and found four frequencies in the spectrum of the liquid These frequencies were also observed in the vapour spectrum where they all showed a *P*, *Q*, *R* structure with a strong central *Q* branch These frequencies were confirmed by Fox & Martin (1938) and are included in table 1.

No data are available for the Raman frequencies of dioxane in the vapour, but the Raman frequencies in the liquid have been reported by several authors. The values quoted in table 1 are taken from a summary by Kahovec & Kohlrausch (1937) of all earlier measurements.

SYMMETRY PROPERTIES

The dioxane molecule possesses $3 \times 14 - 6 = 36$ normal modes of vibration which may, to a first approximation, be described as 8 C-H stretching vibrations, 4 H-C-H bending vibrations, 12 $>CH_2$ rocking and twisting vibrations, 6 C-O and C-C stretching vibrations and 6 C-O-C and O-C-C bending vibrations The frequency ranges in which these various types of vibration may be expected to appear is

TABLE 1. INFRA-RED AND RAMAN FREQUENCIES OF 1·4-DIOXANE

infra-red (cm $^{-1}$)		Raman (cm.$^{-1}$)
vapour	liquid	liquid
		432 (3)
		485 (4)
	711 (0)	
740?	735 (0)	
	753 (0)	
		835 (10)
		852 (0)
880	874 (8)	
889	887 (7)	
	1020 (0*s*)	1014 (6)
P 1042*s*		
Q 1054	1050 (5)	
R 1066		
1087	1086 (7)	
		1111 (3)
1136	1124 (8)	1125 (3)
		1215 (4)
1262	1255 (7)	
1292	1290 (5)	
		1304 (7)
1327*s*	1324 (1*s*)	1331 (0)
1370	1367 (3)	
		1443 (6)
1450	1451 (5)	
1461		1457 (2)
		2663 (0)
		2719 (2)
2863	2855	2858 (8)
2898	2891	2890 (3)
2921	2917	
2970	2961	2967 (9)
		3074 (1)

s denotes unresolved shoulder

Figures in parentheses denote relative intensities of bands.

known from a study of simpler molecules The symmetry classes of the normal vibrations may be derived by the methods of group theory for the various symmetries of the dioxane ring These are given in table 2 together with the infra-red and Raman selection rules and the polarization properties of the Raman lines.

The different structures may now be distinguished by the different numbers of fundamentals allowed in the infra-red and Raman spectra These are summarized in table 3 and compared with the numbers of observed frequencies

It is seen that the less symmetrical structures could only be permitted by postulating that several of the frequencies had not been observed In view of the care taken to find weak bands in the infra-red spectrum, this explanation is unlikely Moreover, only occasional coincidences are observed between the experimental infra-red and Raman frequencies indicating that the selection rules are mutually

TABLE 2. SYMMETRY CLASSES OF NORMAL VIBRATIONS

type of vibration and frequency range (cm $^{-1}$)	irregular	unsym. boat, C_s		sym boat, C_{2v}				chair, C_{2h}				planar, D_{2h}							
		A	B	A_1	A_2	B_1	B_2	A_g	A_u	B_g	B_u	A_g	A_u	B_{1g}	B_{1u}	B_{2g}	B_{2u}	B_{3g}	B_{3u}
8 C-H stretching 2800 to 3000	8	4	4	2	2	2	2	2	2	2	2	1	1	1	1	1	1	1	1
4 H-C-H bending 1350 to 1500	4	2	2	1	1	1	1	1	1	1	1	1	0	1	0	0	1	0	1
12 $>CH_2$ rocking and twisting 750 to 1350	12	6	6	3	3	3	3	3	3	3	3	1	2	1	2	2	1	2	1
6 C-C and C-O stretching 750 to 1350	6	3	3	2	1	2	1	2	2	1	1	2	0	1	0	0	1	0	2
6 O-C-C and C-O-C bending <750	6	4	2	2	2	0	2	2	1	1	2	1	1	1	1	0	1	1	0
selection rules	*IR*	*IR*	*IR*	*IR*	—	*IR*	*IR*	—	*IR*	—	*IR*	—	—	—	*IR*	—	*IR*	—	*IR*
	R	*R*	*R*	*R*	*R*	*R*	*R*	*R*	—	*R*	—	*R*	—	*R*	—	*R*	—	*R*	—
polarization of Raman lines		*P*	*D*	*P*	*D*	*D*	*D*	*P*		*D*		*P*		*D*		*D*		*D*	

IR, infra-red active, *R*, Raman active, *D*, depolarized line, *P*, polarized line

TABLE 3 NUMBERS OF ACTIVE FUNDAMENTALS

	2800 to 3000 cm $^{-1}$		1350 to 1500 cm.$^{-1}$		750 to 1350 cm $^{-1}$		<750 cm $^{-1}$	
	IR	*R*	*IR*	*R*	*IR*	*R*	*IR*	*R*
irregular	8	8	4	4	18	18	6	6
unsym boat	8	8	4	4	18	18	6	6
sym. boat	6	8	3	4	14	18	4	6
chair	4	4	2	2	9	9	3	3
planar	3	4	2	2	7	9	2	3
observed	4	3	2	2	8	8	—	2

exclusive and that the molecule possesses a centre of symmetry Hence it is necessary only to distinguish between the 'chair' and planar configurations The numbers of Raman active fundamentals are the same for the two structures and hence fail to differentiate between them. The structures could be distinguished by the different polarization properties of the Raman lines, but no satisfactory measurements are available The structures may be distinguished, however, by the numbers of fundamentals active in the infra-red spectrum In the region 2800 to 3000 cm.$^{-1}$ four well-defined bands have been observed in the spectrum of the vapour in agreement with the number allowed by the 'chair' structure The two frequencies observed between 1350 and 1500 cm $^{-1}$ agree with the number allowed by both the 'chair' and planar configurations and hence do not distinguish between them. In the region 750 to 1350 cm $^{-1}$ of the vapour spectrum eight separate vibrational frequencies have been observed, though one of these frequencies, viz 1327 cm $^{-1}$, appeared only as a weak unresolved shoulder Two additional bands, viz 1020 and 753 cm $^{-1}$, appeared in the spectrum of the liquid, but these may be due to violation of the selection rules which is well known in the liquid state. Indeed, the former frequency may correspond to the strong Raman frequency at 1014 cm $^{-1}$. Hence the number of frequencies observed in this region does not distinguish conclusively between the two structures The lack of experimental data below 750 cm $^{-1}$ prevents a comparison in this region

Thus although the numbers of infra-red and Raman-active fundamentals are not consistent with structures of low symmetry, unless present in small amounts which escape spectroscopic detection, the evidence does not differentiate conclusively between the 'chair' and planar configurations. The appearance of four well-defined infra-red bands in the region 2800 to 3000 cm $^{-1}$, however, favours the 'chair' structure.

Valency force treatment

The vibration equations for the 'chair' configuration of the dioxane molecule have been obtained using the vector method developed by Eliashevich (1940). The dioxane molecule was treated as a system of six mass points, i, j, k, l, m, n, situated at the vertices of a puckered hexagon (figure 3) All the bond angles γ were assumed to be tetrahedral (109° 28′) in agreement with the electron diffraction data Applying a simple valency force field to the molecule, the potential energy V was expressed by the equation

$$V = \tfrac{1}{2}\sum_{ij} k_{ij} q_{ij}^2 + \tfrac{1}{2}\sum_{i} c_i \alpha_i^2,$$

where q_{ij} = change in the bond length s_{ij} and α_i = change in the bond angle γ_i

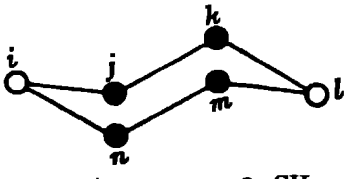


Figure 3. m_a, m_b, masses of the O atom and CH_2 group respectively s_a, s_b, lengths of the C-O and C-C bonds k_a, k_b, stretching constants of the C-O and C-C bonds. c_a, c_b, deformation constants of the C-O-C and O-C-C bond angles

The equations of motion, viz q_{ij} and α_i, were then derived by the methods of Eliashevich (1940) Substituting

$$q_{ij} = (q_{ij})_0 e^{i\nu t} \quad \text{and} \quad \alpha_i = (\alpha_i)_0 e^{i\nu t},$$

equations in ν^2 were obtained whose roots determined the normal vibration frequencies of the molecule These equations were considerably simplified by a knowledge of the symmetry co-ordinates of the normal vibrations which were derived by the methods of group theory as follows

The symmetry operations which leave the 'chair' configuration of the dioxane molecule unchanged are

(*a*) The identity operation, I.

(*b*) Rotation of $\pm\pi$ about the twofold axis, $C_2(z)$.

(*c*) Reflexion across the plane of symmetry perpendicular to the twofold axis, $\sigma_h(xy)$.

(*d*) Inversion about the centre, i.

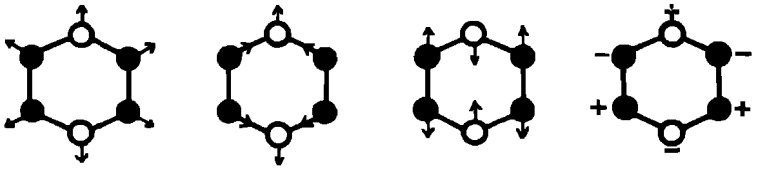
The symmetry properties of the various classes of vibration with respect to these symmetry operations are given in table 4.

TABLE 4. SYMMETRY PROPERTIES OF NORMAL VIBRATIONS

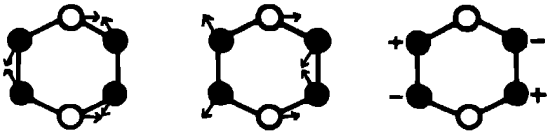
class	I	$c_2(z)$	$\sigma_h(xy)$	ι	selection rules
A_g	+	+	+	+	R
A_u	+	+	−	−	IR
B_g	+	−	−	+	R
B_u	+	−	+	−	IR

+ symmetrical; − antisymmetrical.

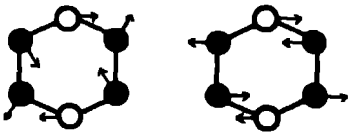
The geometric forms of the normal vibrations were obtained using these symmetry relations and the property that all the normal vibrations must be mutually orthogonal These are given in figure 4



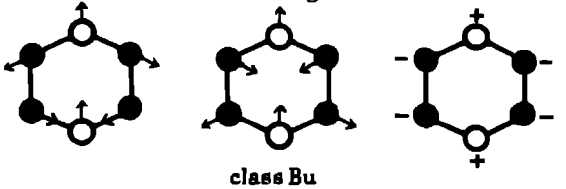
class Ag



class Au



class Bg



class Bu

FIGURE 4

The symmetry co-ordinates were obtained by expressing these vibrations in terms of q and α and are given in figure 5.

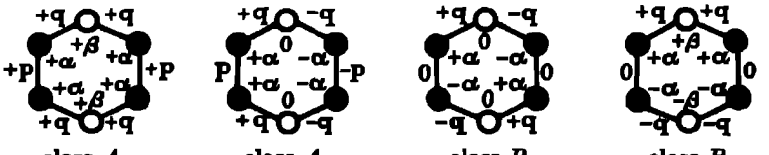


FIGURE 5

By substituting these symmetry co-ordinates into the equations for q_{ij} and α_i the vibration equations were obtained. It is convenient here to introduce the following abbreviations:

$$c = \cos\gamma, \quad s = \sin\gamma, \quad M = \frac{1}{m_a} + \frac{1}{m_b},$$

$$L_1 = \frac{1}{m_a s_a^2}, \quad L_2 = \frac{1}{m_b s_a^2}, \quad L_3 = \frac{1}{m_b s_b^2}, \quad L_4 = \frac{1}{m_b s_a s_b},$$

$$M_1 = 2s_a^2 L_2, \quad M_2 = \tfrac{1}{2}cM_1, \quad M_3 = -sL_1 s_a,$$

$$M_4 = -sL_2 s_a, \quad M_5 = M + cL_1 s_a^2, \quad M_6 = M - cL_1 s_a^2,$$

$$M_7 = -\tfrac{1}{2}s(L_1 + L_4)s_a, \quad M_8 = \tfrac{1}{2}s(L_1 - L_4)s_a, \quad M_9 = M_7 - sL_4 s_a,$$

$$M_{10} = M_8 - sL_4 s_a, \quad M_{11} = 2(1-c)L_1 + 2L_2,$$

$$M_{12} = (2L_1 + L_2 + L_3) + \tfrac{1}{2}c(L_1 - 2L_4),$$

$$M_{13} = -\tfrac{1}{2}(L_1 + L_2 + 2L_4 - cL_1), \quad M_{14} = M_{12} - (2+c)L_1,$$

$$M_{15} = M_{13} + (2+c)L_4, \quad M_{16} = M_{14} + 2(L_3 - cL_4), \quad M_{17} = M_{16} + (2+c)L_1.$$

The vibration equations may then be expressed in the determinant form as follows

Class A_g

$$\begin{vmatrix} (k_a M_5 - \nu^2), & k_b M_2, & c_a M_3, & c_b M_8 \\ 2k_a M_2, & (k_b M_1 - \nu^2), & -c_a M_4, & 2c_b M_4 \\ 2k_a M_3, & -k_b M_4, & (c_a M_{11} - \nu^2), & 2c_b M_{15} \\ k_a M_8, & k_b M_4, & c_a M_{15}, & (c_b M_{12} - \nu^2) \end{vmatrix} = 0$$

Class A_u

$$\begin{vmatrix} (k_a M_6 - \nu^2), & k_b M_2, & c_b M_7 \\ 2k_a M_2, & (k_b M_1 - \nu^2), & 2c_b M_4 \\ k_a M_7, & k_b M_4, & (c_b M_{14} - \nu^2) \end{vmatrix} = 0.$$

Class B_g

$$\begin{vmatrix} (k_a M_6 - \nu^2), & c_b M_9 \\ k_a M_9, & (c_b M_{16} - \nu^2) \end{vmatrix} = 0.$$

Class B_u

$$\begin{vmatrix} (k_a M_5 - \nu^2), & c_a M_3, & c_b M_{10} \\ 2k_a M_3, & (c_a M_{11} - \nu^2), & 2c_b M_{13} \\ k_a M_{10}, & c_a M_{13}, & (c_b M_{17} - \nu^2) \end{vmatrix} = 0.$$

These equations were checked by putting $m_a = m_b = m$, $s_a = s_b = s$, $k_a = k_b = k$ and $c_a = c_b = c$, when it was found that the equations reduced to the earlier equations for cyclohexane (Ramsay & Sutherland 1947) in the following manner

dioxane		cyclohexane
A_g	⟶	A_{1g}, E_{1g}
A_u	⟶	A_{1u}, E_{1u}
B_g	⟶	A_{2g}, E_{1g}
B_u	⟶	A_{2u}, E_{1u}

Assignment of frequencies

Before attempting to evaluate the vibration equations it is helpful to consider the evidence which can be derived from the states of polarization of the Raman lines, the contours of the infra-red bands and the corresponding assignments for cyclohexane Unfortunately, no satisfactory polarization measurements are available to assist the assignment of the Raman frequencies, but it is almost certain that the intense Raman line at 835 cm $^{-1}$ is the totally symmetrical class A_g 'breathing' frequency (cf the strong cyclohexane frequency at 802 cm.$^{-1}$). Since the dioxane frequency is slightly greater than the cyclohexane frequency despite an increase in the mass of the molecule, this would imply that the stretching constant of the C-O bond is greater than the C-C value for cyclohexane The doubly degenerate class E_{1g} Raman line at 426 cm.$^{-1}$ in cyclohexane is split into two component lines at 430 and 454 cm $^{-1}$ in the Raman spectrum of tetrahydropyran (Kahovec & Kohlrausch 1937) This splitting in the case of dioxane probably corresponds to the observed Raman lines at 432 and 485 cm.$^{-1}$ Again, the increase in frequency from cyclohexane to dioxane suggests that the deformation constant for the C-O-C bond angle is somewhat greater than the C-C-C value for cyclohexane Since the class E_{1g} vibrations for cyclohexane split into class A_g and class B_g vibrations for dioxane, and, moreover, since the latter vibrations are not dependent on the value of the C-O-C deformation constant, the 432 cm $^{-1}$ frequency may be assigned to class B_g and the 485 cm. $^{-1}$ frequency to class A_g

The contours of the infra-red bands may be derived from a knowledge of the principal moments of inertia of the dioxane molecule Taking the lengths of the C-C, C-O and C-H bonds as 1·54, 1·46 and 1·09 A respectively and all the bond angles as tetrahedral in agreement with the electron diffraction data (Sutton & Brockway 1937), the principal moments of inertia of the 'chair' structure are

$$I_A = 172 \times 10^{-40} \text{g cm.}^2, \quad I_B = 183 \times 10^{-40} \text{g.cm}^2, \quad I_C = 314 \times 10^{-40} \text{g cm}^2.$$

The molecule thus approximates to a symmetric top. According to Gerhard & Dennison (1933), vibrations involving changes of electric moment parallel to the greatest axis of inertia should give bands with P, Q and R branches Using the mean value of I_A and I_B for the least moment of inertia, the P-R separation was calculated using the doublet formula of these authors and found to be 24 cm $^{-1}$ at 30° C. The contours of the perpendicular bands were obtained by reference to the curves of Gerhard & Dennison (1933) for different values of $\beta = A/C - 1$. For dioxane $\beta = -0{\cdot}436$, hence the perpendicular bands should also show a P, Q, R structure, the Q branch, however, being less prominent than for the parallel bands and the doublet separation being somewhat smaller

The four bands in the vapour spectrum between 2800 and 3000 cm.$^{-1}$ observed by McKinney *et al.* (1937), using a prism grating spectrometer, all showed P, Q, R structures with prominent Q branches Since these bands may be assigned to C-H stretching vibrations, they would be expected to show predominantly parallel type

contours. The P-R separations as measured from the curves of the above authors varied from 18 to 23 cm.$^{-1}$ at 26·6° C in reasonable agreement with the calculated value The three ring vibrations between 750 and 1350 cm $^{-1}$ in the infra-red spectrum should all give perpendicular type bands. Of the six CH_2 rocking and twisting vibrations active in this region, four should give perpendicular type bands and the other two should give bands with hybrid parallel-perpendicular type contours Hence in all there should be seven perpendicular type bands and two with hybrid contours The only band in this region with a P, Q, R structure has well-defined Q and R peaks at 1054 and 1066 cm.$^{-1}$, the P branch appearing as a weak shoulder at 1042 cm.$^{-1}$. The P-R separation agrees with the value calculated for a parallel band, hence this frequency cannot be assigned to a ring vibration It must be due to a $\rangle CH_2$ rocking motion

With this preliminary information a complete assignment of the ring frequencies was attempted by evaluating the vibration equations. Initially it was assumed that the stretching constant of the C-C bond and the deformation constant of the carbon bond angle were the same as for cyclohexane, viz. $k_{C\text{-}C} = 3\,7 \times 10^5$ dynes/cm. and $c_{O\text{-}\widehat{C\text{-}C}} = 1{\cdot}04 \times 10^{-11}$ dyne cm /radian The C-O stretching constant was taken as $4\,45 \times 10^5$ dynes/cm and the deformation constant for the C-O-C bond angle as $1\,4 \times 10^{-11}$ dyne cm./radian, the latter value being chosen to get agreement with the assignment of the 485 cm $^{-1}$ frequency to class A_g The vibration equations were then solved by means of the electrical calculating machine described by Mallock (1933), the frequencies being given in table 5.

TABLE 5 CALCULATED AND OBSERVED RING FREQUENCIES FOR DIOXANE

class	activity	calculated frequency (cm $^{-1}$)	observed frequency (cm $^{-1}$)	percentage difference
A_g	Raman	1111	1111	0·0
		834	835	−0·1
		488	485	+0·6
		390	—	—
A_u	infra-red	1141	1136	+0 4
		927	889 (or 880)	+4 3
		205	—	—
B_g	Raman	1126	1125	+0 1
		429	432	−0·7
B_u	infra-red	982	1020	−3 7
		742	740	+0 3
		233	—	—

The six calculated Raman frequencies were all found to lie remarkably close to observed Raman lines with the exception of the 390 cm $^{-1}$ frequency in class A_g. This frequency agrees well with the corresponding observed frequencies for cyclohexane and tetrahydropyran, viz 384 and 396 cm $^{-1}$ respectively, and is presumably the missing Raman frequency below 700 cm $^{-1}$ permitted by the selection rules.

The calculated values for the infra-red frequencies, however, did not show such good agreement with experiment The calculated frequency at 1141 cm $^{-1}$ lies close to the intense infra-red adsorption band at 1136 cm.$^{-1}$, but the predicted frequencies at 927 and 982 cm.$^{-1}$ do not correspond so closely with observed frequencies. If the weak infra-red frequency at 1020 cm.$^{-1}$ in the spectrum of the liquid is taken as the missing fundamental permitted by the selection rules in the region 750 to 1350 cm $^{-1}$ the most plausible assignment is this weak band to class B_u and the 889 cm $^{-1}$ (or 880 cm $^{-1}$) frequency to class A_u The 1054 cm.$^{-1}$ frequency has been eliminated as a possible ring frequency from contour considerations. The discrepancies between the calculated and observed values for these two frequencies are then of the order of 3 to 4 % and are not unduly large in view of the simplified treatment of the molecule and neglect of the interactions of the hydrogen atoms

The calculated value for the class B_u in-plane bending frequency agrees well with the very weak vapour band observed at 740 cm $^{-1}$. The corresponding frequency for cyclohexane was also found to be exceptionally weak in absorption The two out-of-plane bending frequencies predicted at 205 and 233 cm $^{-1}$ lie beyond the range of prism spectrometers and have not been experimentally observed, but they correspond satisfactorily to a splitting of the doubly degenerate calculated frequency at 206 cm $^{-1}$ for cyclohexane Confirmation of these low frequencies from specific heat data has not been possible since the only existing measurements apply to the liquid state

This completes the assignment of the ring frequencies. Of the twelve frequencies calculated, using four force constants, seven agree with the experimental frequencies to within 1 %, two agree with the observed values to within 3 to 4 %, while the remaining three frequencies have not yet been experimentally observed These three frequencies, however, agree satisfactorily with the corresponding assignments for cyclohexane In view of the simplified force field and treatment of the molecule it is doubtful if any further small changes in the force constants to attempt to improve the agreement between the observed and calculated frequencies could be regarded as significant. The only frequencies below 1200 cm $^{-1}$ which have not been assigned are the 852 and 1014 cm $^{-1}$ Raman lines and the 880 and 1087 cm $^{-1}$ infra-red frequencies. These are presumably hydrogenic frequencies (of the cyclohexane frequencies at 864 and 1015 cm.$^{-1}$).

Discussion of force constants

It is interesting to note that the values of the stretching constant of the C-C bond and deformation constant of the carbon bond angle for dioxane are essentially the same as for cyclohexane. For the latter molecule it was shown that the C-C bond was considerably weaker than in ethane (Ramsay & Sutherland 1947), hence this weakening also persists in dioxane. It is seen from table 6 that the C-C bonds in cyclopropane and ethylene oxide also appear to be weaker than in ethane.

Values for the stretching constants of the C-O bonds in various molecules are given in table 6.

TABLE 6. FORCE CONSTANTS FOR DIOXANE AND RELATED MOLECULES

		stretching constant ($\times 10^5$ dynes/cm)		deformation constant ($\times 10^5$ dynes/cm.)	
author	molecule	C-C	C-O	O-C-C	C-O-C
Kohlrausch (1936)	dimethyl ether	—	4 53	—	0 34
	methyl alcohol	—	4 99	—	—
Bonner (1937)	dimethyl ether	—	4 56	—	0·46
	methyl alcohol	—	5·00	—	—
	ethyl alcohol	3 80	4 48	0·23	—
	ethylene oxide	3 73	4 43	—	—
	cyclopropane	4·04 3·86	—	—	— (0 084)
Stitt (1939)	ethane	4 50	—	—	—
de Hemptinne (1946)	methyl alcohol	—	4 92	—	—
Ramsay & Sutherland (1947)	cyclohexane	3·7	—	0 44	—
present	dioxane	3 7	4·45	0 44	0 66

The C-O stretching constant for dioxane does not differ appreciably from the values obtained for the C-O bonds in dimethyl ether, ethylene oxide and ethyl alcohol The C-O stretching constant for methyl alcohol, however, is somewhat larger The value for this constant obtained by de Hemptinne (1946), using a detailed potential function, agrees with the earlier values of Kohlrausch (1936) and Bonner (1937) obtained by treating the molecule as a diatomic system Hence the high value of the C-O force constant for methyl alcohol probably implies a real strengthening of the C-O bond Little information is available on the deformation constants of oxygen bond angles, but the value for dioxane appears to be considerably larger than for dimethyl ether. While the values of the C-C and C-O stretching constants for dioxane are similar to those found for related molecules, any attempt to interpret the chemical reactivities of these bonds in terms of their force constants should be treated with caution.

The author is indebted to Dr G. B. B M. Sutherland for valuable advice and discussion, to the Director of the University Mathematical Laboratory for permission to use the electrical calculating machine, and to Messrs BX Plastics, Ltd , for financial assistance.

REFERENCES

Böeseken, J , Tellegen, F. & Henriquez, P. C 1935 *Rec. Trav. chim Pays-Bas*, **54**, 733.
Bonner, L G 1937 *J. Chem. Phys* **5**, 293.
Elliashevich, M A. 1940 *C.R. Acad. Sci. U.R.S.S.* **28**, 604.
Fox, J. J & Martin, A. E. 1938 *Proc. Roy. Soc.* A, **167**, 257.
Gerhard, S. L. & Dennison, D. M 1933 *Phys Rev.* **43**, 197.
de Hemptinne, M. 1946 *Trans Faraday Soc.* **42**, 5.

Hunter, E. C E. & Partington, J. R 1933 *J Chem. Soc.* p. 87.
Kahovec, L & Kohlrausch, K W F 1937 *Z Phys. Chem* B, **35**, 29.
Kohlrausch, K W. F. 1936 *Mh Chem.* **68**, 349
Kubo, M 1936 *Sci. Pap Inst. Phys Chem. Res. Tokyo*, **30**, 238.
McKinney, D. S , Leberknight, C. E. & Warner, J C. 1937 *J. Amer. Chem Soc* **59**, 481.
Mallock, R. R. M. 1933 *Proc Roy Soc.* A, **140**, 457
Ramsay, D A. & Sutherland, G B B M 1947 *Proc. Roy Soc.* A, **190**, 245
Sängewald, R & Weissberger, A. 1929 *Phys Z.* **30**, 268.
Schwingel, C. H & Greene, E. W. 1934 *J. Amer. Chem. Soc.* **56**, 653
Smyth, C. P & Walls, W. S 1931 *J Amer. Chem Soc* **53**, 2115.
Stitt, F. 1939 *J. Chem. Phys* **7**, 297.
Sutton, L E & Brockway, L O 1935 *J. Amer. Chem. Soc.* **57**, 473
Vaughan, W. C 1939 *Phil Mag* **27**, 669.
Williams, J. W 1930 *J Amer Chem. Soc.* **52**, 1831.

INDEX TO VOLUME 190 (A)

Accuracy of atomic co-ordinates derived from Fourier series in X-ray structure analysis. III. (Booth), 482

Accuracy of atomic co-ordinates derived from Fourier series in X-ray structure analysis IV. The two-dimensional projection of oxalic acid (Booth), 490.

Accuracy of bond length estimations obtained by double Fourier series methods from X-ray data (Robertson & White), 329.

Adair, G S. & Adair, M E The density increments of proteins, 341.

Adair, M. E *See* Adair & Adair

Adsorption of vapours on mercury III Polar substances (Kemball), 117

Aluminium-manganese-zinc Intermetallic compound phases (Raynor & Wakeman), 82.

Assignment of the slow-neutron-produced activities of thallium and the dual disintegration of radium E (Broda & Feather), 20.

Atmospheric tides in the ionosphere. II Lunar tidal variations in the *F* region near the magnetic equator (Martyn), 273

Banerjea, B K. On the propagation of electromagnetic waves through the atmosphere, 67

Batchelor, G. K & Townsend, A A Decay of vorticity in isotropic turbulence, 534

Beevers, C A. & Cochran, W The crystal structure of sucrose sodium bromide dihydrate, 257.

Booth, A D The accuracy of atomic co-ordinates derived from Fourier series in X-ray structure analysis. III, 482

Booth, A. D. The accuracy of atomic co-ordinates derived from Fourier series in X-ray structure analysis IV. The two-dimensional projection of oxalic acid, 490

Born, M. & Green, H. S. A general kinetic theory of liquids III Dynamical properties, 455

Bragg, Sir Lawrence & Nye, J F. A dynamical model of a crystal structure, 474

Broadbent, D & Jánossy, L Local and extensive penetrating cosmic ray showers, 497.

Broda, E & Feather, N. The assignment of the slow-neutron-produced activities of thallium and the dual disintegration of radium E, 20

Brookfield, K J , Fitzpatrick, H D N , Jackson, J F , Matthews, J B & Moelwyn-Hughes, E A The escape of molecules from a plane surface into a still atmosphere, 59.

Catalytic dehydrogenation of naphthenes I Kinetic study (Herington & Rideal), 289

Catalytic dehydrogenation of naphthenes II. Competition experiments and energies of C-H bonds (Herington & Rideal), 309

Cochran, W *See* Beevers & Cochran

Coleman, D & Howitt, F O Studies on silk proteins I The properties and constitution of fibroin. The conversion of fibroin into a water-soluble form and its bearing on the phenomenon of denaturation, 145

Control of amplitude in phase-contrast microscopy (Taylor), 422

Cosmic-ray bursts and shower spread under large thicknesses of lead (Juritz & Mohr), 426

Crystal structure (Bragg & Nye), 474

Crystal structure of some molecular complexes of 4 4′-dinitrodiphenyl II. The complexes with benzidine, *NN*:*N′N′*-tetramethylbenzidine, 4 4′-dihydroxydiphenyl, 4-iododiphenyl and 4-bromodiphenyl (Saunder), 508

Crystal structure of sucrose sodium bromide dihydrate (Beevers & Cochran), 257

Cwilong, B. M Sublimation in a Wilson chamber, 137

Cyclohexane. Vibration spectrum and molecular configuration (Ramsay & Sutherland), 245

Davidson, P M The establishment of the dilatation effect in a field of sound waves, 418
Davies, D. R Turbulence and diffusion in the lower atmosphere with particular reference to the lateral effect, 232
Decay of vorticity in isotropic turbulence (Batchelor & Townsend), 534.
Density increments of proteins (Adair & Adair), 341.
Dynamical model of a crystal structure (Bragg & Nye), 474.

Elastic scattering of 6·5 MeV deuterons by deuterium, helium and other light elements (Guggenheimer, Heitler & Powell), 196.
Escape of molecules from a plane surface into a still atmosphere (Brookfield, Fitzpatrick, Jackson, Matthews & Moelwyn-Hughes), 59
Establishment of the dilatation effect in a field of sound waves (Davidson), 418

Fage, A & Sargent, R. F Shock-wave and boundary-layer phenomena near a flat surface, 1
Feather, N *See* Broda & Feather
Fibroin Properties and constitution (Coleman & Howitt), 145.
Fitzpatrick, H. D N *See* Brookfield, Fitzpatrick, Jackson, Matthews & Moelwyn-Hughes
Fox, L Some improvements in the use of relaxation methods for the solution of ordinary and partial differential equations, 31

General kinetic theory of liquids III Dynamical properties (Born & Green), 455.
Gowan, E. H Night cooling of the ozonosphere, 227
Gowan, E. H Ozonosphere temperatures under radiation equilibrium, 219.
Green, H S *See* Born & Green
Guggenheimer, K M, Heitler, H & Powell, C F The elastic scattering of 6·5 MeV deuterons by deuterium, helium and other light elements, 196

Heitler, H., May, A N. & Powell, C F The scattering of 4 2 MeV protons by deuterium helium and other light elements, 180
Heitler, H *See* Guggenheimer, Heitler & Powell
Herington, E F G. & Rideal, E K On the catalytic dehydrogenation of naphthenes. I Kinetic study, 289
Herington, E F G. & Rideal, E K The catalytic dehydrogenation of naphthenes II Competition experiments and energies of C-H bonds, 300
Howitt, F O *See* Coleman & Howitt
Huang, K X-ray reflexions from dilute solid solutions, 102.

Influence of space charge on thermionic emission velocities (Klemperer), 376.
Intermetallic compound phases of the system aluminium-manganese-zinc (Raynor & Wakeman), 82.
Ionosphere. Atmospheric tides (Martyn), 273

Jackson, J F *See* Brookfield, Fitzpatrick, Jackson, Matthews & Moelwyn-Hughes
James, R. W. & Saunder, D H Some apparent periodic errors in the crystal lattice of the molecular complexes of 4:4′-dinitrodiphenyl with 4-iodo- and 4-bromodiphenyl, 518
Jánossy, L *See* Broadbent & Jánossy
Juritz, J W F & Mohr, C B. O Cosmic-ray bursts and shower spread under large thicknesses of lead, 426

Kemball, C The adsorption of vapours on mercury III. Polar substances, 117
Klemperer, O. Influence of space charge on thermionic emission velocities, 376

Local and extensive penetrating cosmic ray showers (Broadbent & Jánossy), 497
Lunar tidal variations in the *F* region near the magnetic equator. Atmospheric tides in the ionosphere II (Martyn), 273

McCready, L. L., Pawsey, J. L & Payne-Scott, R Solar radiation at radio frequencies and its relation to sunspots, 357
Martyn, D F Atmospheric tides in the ionosphere II Lunar tidal variations in the *F* region near the magnetic equator, 273.
Matthews, J B *See* Brookfield, Fitzpatrick, Jackson, Matthews & Moelwyn-Hughes
May, A N & Powell, C. F The scattering of 4 2 MeV protons by protons, 170
May, A. N *See* Heitler, May & Powell
Measurements in the turbulent wake of a cylinder (Townsend), 551
Moelwyn-Hughes, E. A *See* Brookfield, Fitzpatrick, Jackson, Matthews & Moelwyn-Hughes
Mohr, C B O *See* Juritz & Mohr

Naphthenes Catalytic dehydrogenation I Kinetic study (Herington & Rideal), 289
Naphthenes Catalytic dehydrogenation II Competition experiments (Herington & Rideal), 309
Night cooling of the ozonosphere (Gowan), 227
Nye, J F *See* Bragg & Nye.

Oxide cathodes The effects of the coating-core interface on conductivity and emission (Wright), 394
Ozonosphere. Night cooling (Gowan), 227
Ozonosphere temperatures under radiation equilibrium (Gowan), 219.

Pawsey, J L. *See* McCready, Pawsey & Payne-Scott
Payne-Scott, R *See* McCready, Pawsey & Payne-Scott
Powell, C. F *See* May & Powell, Heitler, May & Powell *and* Guggenheimer, Heitler & Powell.
Propagation of electromagnetic waves through the atmosphere (Banerjea), 67
Proteins (Adair & Adair), 341.

Radium E Dual disintegration (Broda & Feather), 20
Ramsay, D A The vibration spectrum and molecular configuration of 1·4-dioxane, 562.
Ramsay, D A & Sutherland, G B B M The vibration spectrum and molecular configuration of cyclohexane, 245.
Raynor, G V & Wakeman, D. W The intermetallic compound phases of the system aluminium-manganese-zinc, 82.
Rideal, E K *See* Herington & Rideal
Robertson, J. M. & White, J G. The accuracy of bond-length estimations obtained by double Fourier series methods from X-ray data, 329

Sargent, R F *See* Fage & Sargent
Saunder, D. H. The crystal structure of some molecular complexes of 4 4′-dinitrodiphenyl. II. The complexes with benzidine, *NN N′N′*-tetramethylbenzidine, 4·4′-dihydroxydiphenyl, 4-iododiphenyl and 4-bromodiphenyl, 508.

Saunder, D H *See* James & Saunder.
Scattering of 4·2 MeV protons by deuterium, helium and other light elements (Heitler, May & Powell), 180.
Scattering of 4·2 MeV protons by protons (May & Powell), 170.
Shock-wave and boundary-layer phenomena near a flat surface (Fage & Sargent), 1.
Silk proteins. I (Coleman & Howitt), 145.
Solar radiation at radio frequencies and its relation to sunspots (McCready, Pawsey & Payne-Scott), 357.
Some apparent periodic errors in the crystal lattice of the molecular complexes of 4.4'-dinitrodiphenyl with 4-iodo- and 4-bromodiphenyl (James & Saunder), 518
Some improvements in the use of relaxation methods for the solution of ordinary and partial differential equations (Fox), 31.
Sondheimer, E. H. & Wilson, A H. The theory of the magneto-resistance effects in metals, 435.
Studies on silk proteins I The properties and construction of fibroin. The conversion of fibroin into a water-soluble form and its bearing on the phenomenon of denaturation (Coleman & Howitt), 145
Sublimation in a Wilson chamber (Cwilong), 137.
Sutherland, G. B. B M. *See* Ramsay & Sutherland.

Taylor, E W The control of amplitude in phase-contrast microscopy, 422.
Thallium Slow-neutron-produced activities (Broda & Feather), 20.
Theory of the magneto-resistance effects in metals (Sondheimer & Wilson), 435.
Townsend, A A Measurements in the turbulent wake of a cylinder, 551.
Townsend, A A. *See* Batchelor & Townsend
Turbulence and diffusion in the lower atmosphere with particular reference to the lateral effect (Davies), 232

Vibration spectrum and molecular configuration of cyclohexane (Ramsay & Sutherland), 245.
Vibration spectrum and molecular configuration of 1.4-dioxane (Ramsay), 562.

Wakeman, D W. *See* Raynor & Wakeman
White, J G. *See* Robertson & White
Wilson, A. H *See* Sondheimer & Wilson
Wright, D A Oxide cathodes. The effect of the coating-core interface on conductivity and emission, 394

X-ray reflexions from dilute solid solutions (Huang), 102.